Functional characterisation of *Myosin V and Broad Complex* during *Drosophila* oogenesis

Identification and analysis of myosin genes in *Drosophila melanogaster*

George Iltchev Tzolovsky

Thesis presented for the degree of Doctor of Philosophy
Institute of Cell and Molecular Biology
The University of Edinburgh
2002

Abstract

Myosins constitute a superfamily of motor proteins that convert energy from ATP hydrolysis into mechanical movement along actin filaments. Phylogenetic analysis currently places myosins into 18 classes based on class specific features of their conserved motor domain.

To find how many myosin genes *Drosophila* has, we have analysed the *Drosophila* Genome Sequence. We identified 13 myosin genes. Phylogenetic analysis based on the sequence comparison of the myosin motor domains, as well as the presence of class specific domains, suggests that *Drosophila* myosins can be divided into nine major classes. Myosins belonging to previously described classes I, II, III, V, VI and VII are present. Molecular and phylogenetic analysis indicates that the fruitfly genome contains at least five new myosins. Three of them fall into previously described myosin classes, I, VII and XV. Another myosin is a homologue of the mouse and human PDZ-containing myosins, forming the recently defined class XVIII myosins. The fifth myosin shows a unique domain composition and a low homology to any of the existing classes. We propose that this is classified when similar myosins are identified in other species.

Myosin V is one of the best studied unconventional myosin in vertebrates. It has been implicated in a number of functions such as vesicular transport and polarised cell growth. *didum* which encodes Myosin V in *Drosophila*, is strongly expressed during oogenesis. RNA analysis revealed that the gene produces at least three alternative splice forms expressed differentially in the different tissues and developmental stages. We have generated antibodies to analyse the protein distribution. During the early stages of oogenesis its expression pattern localises to the nurse cells delivering materials (RNA and proteins) to the developing oocyte. Later (stage10B-13) myosin V is expressed in the cytoplasm of all follicle. Additionally we have been able to identify two null mutants for *didum*. The mutants die as first instar larvae due to severely delayed development. The mutant larvae are unable to move showing a general larval system failure.

Vertebrate Myosin V was shown to be expressed mainly in the neurons participating in the intracellular transport of vesicles. It is possible that mispositioning of the vesicles, due to myosin V deficiency could affect the developing CNS in *didum* mutants and contribute to the observed defects.

The *Broad-Complex (BR-C)*, an early ecdysone response gene that functions during metamorphosis and which encodes a family of zinc-finger proteins, is expressed in a dynamic pattern during oogenesis. Its late expression in the lateral-dorsal-anterior follicle cells is related to the morphogenesis of the chorionic appendages. All four zinc finger isoforms are expressed in oogenesis, which is consistent with the abnormal appendage phenotypes resulting from their ectopic expression. We investigated the mechanism by which the *BR-C* affects chorion deposition by using BrdU to follow the effects of *BR-C* misexpression on DNA replication and *in situ* hybridisation to ovarian RNA to evaluate chorion gene expression. Ectopic *BR-C* expression leads to prolonged endoreplication and to additional amplification of genes, besides the chorion genes, at other sites in the genome. The pattern of chorion gene expression is not affected along the anterior-posterior axis, but the follicle cells at the anterior of the oocyte fail to migrate correctly in an anterior direction when *BR-C* is misexpressed. We conclude that the target genes of the *BR-C* in oogenesis include a protein essential for endoreplication and chorion gene amplification. This may provide a link between steroid hormones and the control of DNA replication during oogenesis.

Acknowledgements

I would like to thank my supervisor, Professor Mary Bownes for her help and support during the years of my studies. I would also like to thank all members of the Bownes group, past and present: Debiao, Wu-Min, Tim, Stephen, Kevin, Sheng-yin, Leeanne, Hadas, Sofia, Jun, Claudia, Paolo, and Dot.

Special thank to Kathleen for her help and patience and Sheila Milne for typing our papers.

I would like to thank Professor Sir Ken Murray and the Darwin Trust of Edinburgh for allowing me this opportunity to come and study in Edinburgh.

Thanks to all friends and flatmates who helped make my time in Edinburgh so enjoyable: Lena, Uzma, Milly, Tshering, Thorsten, Sabrina, George, Andrea, Hazel, Axel, Vivian, Kampanad and many more. Thanks to all friends I met at the salsa scene in Edinburgh and the members of our karate club.

Finally I would like to thank my family for their support and encouragement at all times.

Abbreviations

General Abbreviations:

λ	lambda
Δ	delta
♂	male
β	beta
μ	micro- (1.10^{-6})
♀	female
[x]	Concentration X
^{32}P	β-emitting isotope of phosphorus
aa	amino acids
A_{xnm}	Absorption at X nm
bp	base pair(s)
C-	Carboxy - (terminal)
cDNA	complementary
cm	centimetre(s)
C^{o}	degrees centigrade
g	gram
g	G centrifugal force
kb	kilobase(s)
krpm	kilo (1000) revolutions per minute
l	litre
M	Molar
ml	millilitre (1.10^{-3})
mM	millimolar (1.10^{-3})
mol	-moles
n	nano- (1.10^{-9})
N-	Amino - (terminal)

nm	nanometres (1.10^{-9}m)
nt	nucleotides
O/N	overnight
OD_{xnm}	Optical Density at X nm
p	pico- (1.10^{-12})
PCR	Polymerase Chain Reaction
pers. comm.	personal communications
pH	-log10 [hydrogen ion]
RT-PCR	Reverse Transcription PCR
U	Units
UTR	Untranslated region
UV	Ultraviolet
v/v	volume to volume ratio
w/v	weight to volume ratio
w/w	weight to weight ratio

Chemicals:

Amp	Ampicillin
AP	Alkaline phosphatase
$CaCl_2$	Calcium chloride
$CHCl_3$	Chloroform
CIAP	Calf intestinal alkaline phosphatase
dATP	2' deoxyadenosine-5'-triphosphate
dCTP	2' deoxycytosine-5'-triphosphate
ddH2O	double distilled water
dGTP	2' deoxyguanosine-5'-triphosphate
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid

DNase	Deoxyribonuclease
dNTP	deoxynucleotide-5'-triphosphate
dTTP	2' deoxythymidine-5'-triphosphate
dUTP	2' deoxyuridine-5'-triphosphate
EtBr	Ethidium bromide
EtOH	Ethanol
HCl	Hydrochloric acid
Hepes	N-[2-hydroxyethyl] piperazine-N'-[2-ethanesulphonic acid]
Kan	Kanamycin
KCl	Potassium chloride
KH_2PO_4	Potassium dihydrogen phosphate
KOAc	Potassium Acetate
KOH	Potassium hydroxide
MeOH	Methanol
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulphate
$MnCl_2$	Manganese chloride
mRNA	messenger RNA
Na_2HPO_4	Disodium hydrogen phosphate
NaAc	Sodium acetate
NaCl	Sodium chloride
NaH_2PO_4	Sodium dihydrogen phosphate
NaOH	Sodium hydroxide
NH_4OAc	Ammonium acetate
PBS	Phosphate buffered saline
PIPES	Piperazine-N,N'-bis[2-ethanesulphonic acid]
RbCl	Rubidium chloride
RNA	Ribonucleic acid
rRNA	ribosomal RNA
SDS	Sodium dodecyl sulphate

Tris	Tris(hydroxymethyl)-amino-methane
Triton X-100	Octylphenoxypolyethoxyethanol
tRNA	transfer RNA
Tween-20	Polyoxyethylene sorbitan monolaurate

Single letter amino acid code:

A	Alanine	L	Leucine
R	Arginine	K	Lysine
N	Asparagine	M	Methionine
D	Aspartate	F	Phenylalanine
C	Cysteine	P	Proline
Q	Glutamine	S	Serine
E	Glutamate	T	Threonine
G	Glycine	W	Tryptophan
H	Histidine	Y	Tyrosine
I	Isoleucine	V	Valine

Table of Contents

Chapter One: Introduction

Drosophila melanogaster is a powerful model organism used for research into a number of fundamental aspects of development. A major task is to elucidate the complex mechanisms controlling the development of the oocyte. Much remains to be learned about the mechanisms by which a single cell develops and differentiates into all cell types of the multicellular organism. What mechanisms generate and maintain the cytoplasmic asymmetry within syncytial cysts during oocyte differentiation, what regulates the transfer of macromolecules and organelles from the nurse cells to the oocyte, what triggers the apoptotic events in the nurse cells and subsequently in the follicle cells, what kinds of interactions exist between the germline cells and the surrounding somatic follicle cells, and how is the positional information, essential for the establishment of body pattern, generated? The answers to these questions will provide insights into many aspects of cellular biology.

1.1 Overview of Drosophila oogenesis

1.1.1 Morphology of *Drosophila* oogenesis

An adult *Drosophila* female has a pair of ovaries connected to the uterus via a common oviduct (King 1970; Bownes and Dale 1982). Each ovary is composed of 16-18 parallel strings (ovarioles) of progressively mature egg chambers which subsequently develop into mature eggs in the course of approximately 5 days (Fig.1.1.1A). Two layers of somatic muscle fibres, an outer peritoneal sheath and an inner epithelial sheath overlay the ovaries and assist with the progression of developing egg chambers.

Drosophila oogenesis has been divided into 14 stages (King 1970; Spradling 1993) representing morphologically distinct periods (Fig.1.1.1C). Oogenesis is initiated within the anterior compartment of the ovariole, the germarium. The germline stem cell divides to produce a cystoblast and regenerate a stem cell (Spradling 1993; de Cuevas et al. 1997).

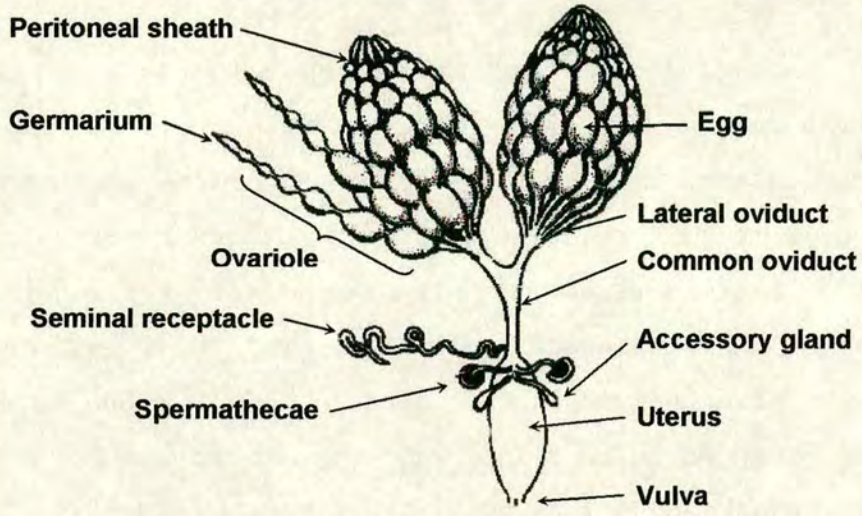


Figure 1.1.1A Reproductive system of adult female *D. melanogaster*

Dorsal view of *Drosophila* ovaries (King 1970). An adult female fly has two ovaries connected to the uterus by a common oviduct. Each ovary consists of approximately 16 ovarioles of progressively maturing egg chambers. At the distal tip of each ovariole there is a distinct region, where oogenesis starts, called the germarium (adapted from Spradling et al. 1997). The mature eggs are fertilised in the uterus from the sperm stored in the seminal receptacle and spermathecae.

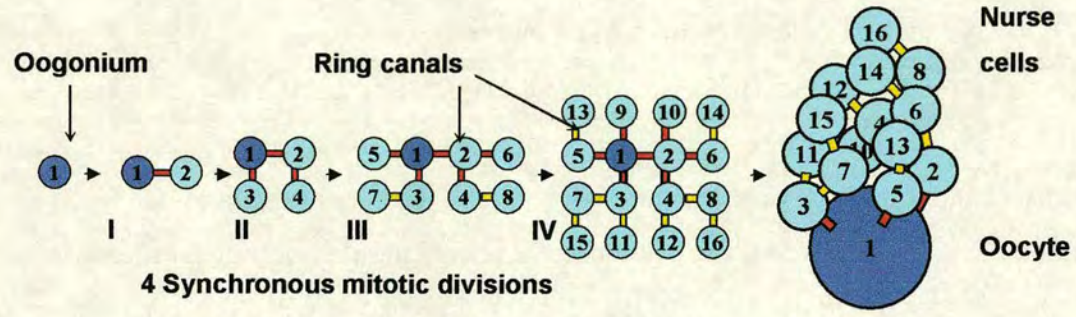


Figure 1.1.1B Pattern of intracellular bridges (ring canals)

The cystoblast (oogonium) undergoes four synchronous rounds of division to produce a 16-cell germ-line cyst (cystocytes). Due to incomplete cytokinesis the daughter cystocytes remain interconnected by specialised cleavage furrows called ring canals. During the mitoses the fusome, a vesicular organelle, always associates with one of the poles of the mitotic spindle and is inherited by one of the daughter cells (labelled in dark blue). The daughter cell with four ring canals and a fusome always becomes the oocyte. The other 15 cells develop and differentiate into nurse cells.

The cystoblast undergoes four synchronous mitotic divisions to yield a cyst of 16 interconnected cells, cystocytes. One of the cells, usually the posterior-most cystocyte with four ring canals develops into the oocyte (Fig. 1.1.1B) (Robinson et al. 1994; Robinson and Cooley 1997). The remaining 15 cells become polyploid nurse cells that provide a variety of organelles, liposomes, and RNAs to the developing oocyte and embryo. The follicle cells also exhibit polyploidism, but to a lesser extent.

Germaria can be subdivided into four cytologically distinct regions (Fig.1.1.1C) (King 1970; Spradling et al. 1997). Region 1 contains germ line cyst cells and the dividing cystoblast. In region 2a lie the newly formed 16-cell cysts, whereas in region 2b the cysts become lens shaped with the proocyte positioned in the centre of the cyst. Somatic stem cells are located at the border between 2a/2b regions and divide to produce profollicle cells. The completed 16-cell cyst acquires a monolayer of follicle cells in region 3 and buds from the germarium as a new egg chamber. The cell complex in region 3 is referred to as a stage 1 egg chamber. At stage 1 the oocyte is located at the posterior pole of the cyst and remains in this position throughout the oogenesis.

From stage 2 to 6 egg chambers increase in size and become more elongated in shape. The oocyte grows slower than the nurse cells during these early stages. From stages 7 through 10 the size of the oocyte increases rapidly compared to the nurse cells as a consequence of the uptake of yolk proteins. Yolk proteins are synthesised in the fat body and follicle cells (Hames and Bownes 1978; Isaac and Bownes 1982). The yolk protein produced in the fat body is secreted in the haemolymph, from where it is selectively endocytosed into the oocyte. The yolk from the follicle cells is directly transported to the oocyte.

The somatically derived follicle cells remain as a single cell layer around the entire egg chamber until stage 8 (King 1970; Dobens and Raftery 2000). At stage 9 the follicle cells start a series of migrations and differentiate into several subtypes (Fig.1.1.1C). Initially the majority of the follicle cells reorganise themselves to form a tightly packed, columnar layer around the oocyte. The remaining follicle cells, approximately 5%, form a thin stretched layer around the nurse cells (nurse cell associated follicle cells).

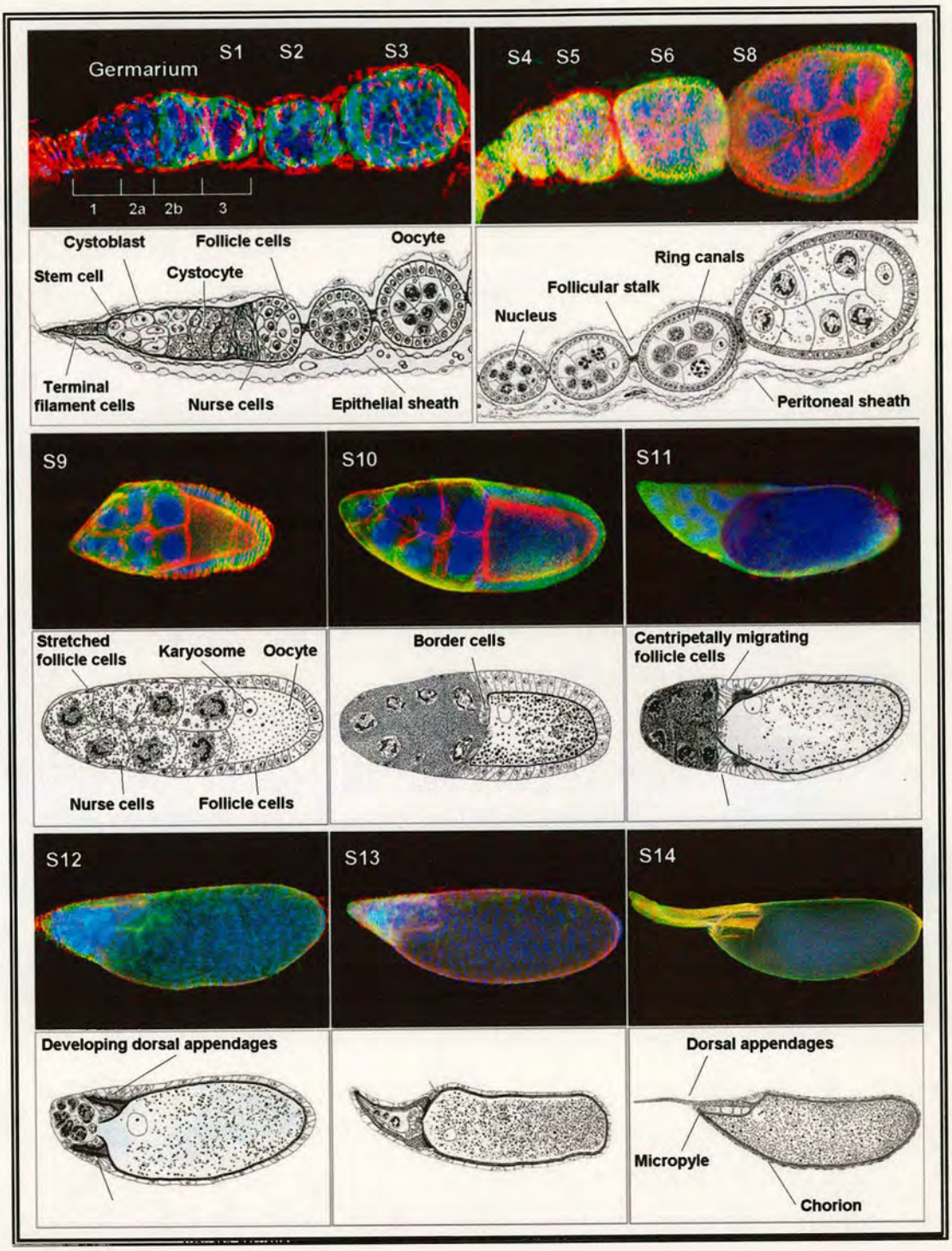
Figure 1.1.1C: Diagram of a single ovariole

Each ovariole usually contains up to seven progressively mature egg chambers, separated by short stalks of specialised follicle cells. Staging based on King 1970. The germarium contains two types of cells: germ-line cells and somatically derived follicle cells. In region 1 a pro-oocyte is formed as a result of asymmetric stem cell division. The cystoblast undergoes four incomplete mitotic divisions to form a syncytial cyst, which flattens as it exits region 2A. As the cyst moves in a posterior direction one of the cystocytes is selected to become oocyte. The remaining 15, germ-line derived cells, become nurse cells. The cyst is enveloped by follicle cells in region 2B to form a stage 1 egg chamber (region 3). In addition various somatic cells are involved in the formation and maintenance of the developing egg chamber: somatic stem cells, terminal filament cells, and inner sheath cells.

At stage 9 the egg chamber consists of three cell types: the oocyte with its nucleus (germinal vesicle), the nurse cells, and the follicle cells. In addition three distinct subpopulations of follicle cells can be recognised: the columnar follicle cells, the nurse cell associated follicle cells, and the border cells (see stage10). At stage 10B the follicle cells begin dumping their cytoplasm into the developing oocyte. This process overlaps with the migration of the centripetal follicle cells to cover the anterior tip of the oocyte (stage 11). As the oocyte develops, the follicle cells stretch to maintain an intact epithelium. Two small groups of follicle cells migrate anteriorly to produce the dorsal appendages. The follicle cells surrounding the oocyte start depositing the eggshell. At stage 14 the mature egg is covered with an eggshell (chorion) containing several external chorionic structures. Two respiratory (dorsal) appendages protrude from the anterior end of wild type *Drosophila* eggs. Just below the appendages is located a structure called an operculum (the larval exit) with a centrally positioned micropyle. On the posterior pole is found an aeropyle, a small region of specialised chorion with respiratory functions in some Drosophila's species.

The actin cytoskeleton is stained with fluorescently labelled phalloidin (Red). The nuclei are stained for DNA with DAPI (Blue). The cells are immunohystochemically stained with antibodies for nonmuscle myosin II (Green). The black and white graphics are reproduced from King 1970.

Figure 1.1.1C: Diagram of a single ovariole



At stage 10A the oocyte occupies the posterior half of the egg chamber, while the 15 nurse cells fill the anterior half of the egg chamber. During stages 9-10 a group of 6-10 anterior follicle cells migrate through the nurse cells to form a cluster of border cells at the anterior surface of the oocyte (Montell et al. 1992; Edwards and Kiehart 1996) (Fig.1.1.1C). Later in oogenesis (stage 11) they will be required for formation of the micropyle, a structure through which the sperm enters the oocyte for fertilisation. At stage 10B the anterior columnar follicle cells move centripetally toward the border cells to cover the anterior end of the oocyte (Keyes and Spradling 1997).

The dorsal appendages are required to deliver oxygen through the eggshell to the developing egg (Spradling 1993). They are specified by two groups of columnar follicle cells from the dorsal-anterior region, which start to migrate anteriorly at stage 11 (Deng and Bownes 1997). The elongation of the dorsal appendages is completed by stage 14. During stages 10B to 12 the nurse cells transfer their cytoplasm through the ring canals to the oocyte and degenerate at the subsequent stages, 13-14. During stages 9 and 10 the follicle cells secrete the vitelline membrane (Burke et al. 1987), which block completely the yolk uptake by stage 11. Then the follicle cells produce the end and exochorion layers to protect the mature egg. Then they too undergo apoptosis. The egg (Fig.1.1.1C) is able to develop to the blastoderm stage utilising the stored maternal mRNAs.

1.1.2 Polarity determination and pattern formation

1.1.2.1 Germline differentiation in early oogenesis

As oogenesis progresses, communication takes place between the germ line cells and the somatic follicle cells in order to determine the polarity and organisation of the developing egg chamber. These cell-cell interactions are critical in the establishment of firstly Anterior/Posterior polarity and subsequently the Dorsal/Ventral polarity of the egg (Spradling 1993; Deng and Bownes 1998). The actin and microtubule cytoskeletons have a central role in establishing and maintaining the polarity in the egg (Theurkauf 1994; Knowles and Cooley 1994). Developing germ line cysts contain a large cytoplasmic organelle called a fusome, which is built mainly of components of the submembrane cytoskeleton (Lin and

Spradling 1995). It was found that membrane cytoskeletal proteins (α-spectrin, β-spectrin, dynein), CyclinA, and Hu-li tai shao (Hts), *Drosophila* homologue of the mammalian adducin, are major fusome constituents (Spradling et al. 1997; Zaccai and Lipshitz 1996). During the first cystoblast division, a vesicular structure called spectrosome is inherited asymmetrically by the daughter cells (Fig. 1.1.1B). That determines which of the two daughter cells will become the oocyte. This in turn leads to the formation of an active MTOC (microtubule organising centre) in that single cell (Fig. 1.1.2.1). During the subsequent divisions the spectrosome continues to segregate asymmetrically, and grows to form a structure called a fusome that spans all the cystocytes by means of ring canals, before disappearing after the fourth division (McKearin 1997; Spradling et al. 1997). The fusome appears to have a crucial role in establishing and polarising the microtubule network. It is thought that the fusome is the "oocyte determination factor". Several genes have been shown to be critical for MTOC specification, such as: *Bicaudal (Bic-D), egaliterian (egl),* and *stonewall (stwl)* (Clark and McKearin 1996). Mutations in these genes produce cysts of 16 nurse cells and no oocyte (Schupbach and Wieschaus 1991)

1.1.2.2 Origin of polarity

The germline cyst is already formed in region 2 of the germarium. Selective transport of mRNAs and proteins to one of the 16 cystocytes, which will develop into the oocyte has been observed. The first sign of A/P polarisation is the pro-oocyte positioning at the posterior pole of the newly formed stage 1 egg chamber (Fig.1.1.2.3) (van Eeden and St Johnston 1999). This movement requires the products of several genes including *armadillo* (Peifer et al. 1993), five *spindle* genes (Morris and Lehmann 1999), and *dicephalic* (Gonzalez-Reyes and St Johnston 1994). It was observed that the follicular epithelium forms a symmetrical pattern along the A/P axes prior to the oocyte movement, with the polar cells located at both anterior and posterior ends (Margolis and Spradling 1995). It was suggested that the polar cells send signals to regulate the movement of the oocyte , and maintain its posterior position (Peifer et al. 1993). The *hedgehog (hh)* signalling pathway appears to be required for the differentiation of pairs of polar follicle cells (Forbes et al. 1993). The product of *hedgehog* has a morphogenetic function and regulates cyst

envelopment in the germarium region 2. The *Notch-Delta* signalling pathway, including neurogenic genes, has been also shown to be involved in the processes of cyst encapsulation and cell differentiation. *Delta* encodes a transmembrane protein which interacts with the product of *Notch.* Muskavitch (1994) has found that *Notch* and *Delta* exhibit a lateral inhibition mechanism in specifying cell fate. Mutations in the genes result in a hyperplasia of stalk cells early in oogenesis and later in the loss of polar cells (Larkin et al. 1996). Some mutations in *Notch* lead to fused egg chambers resulting from the encapsulation of two or more germline cysts (Goode et al. 1996).

1.1.2.3 Anterior/Posterior axis determination

Gurken (grk), a transforming growth factor alpha (TGF-α) homologue signalling molecule has been found to be the key signal in defining the anterior/posterior axis (Gonzalez-Reyes et al. 1995; Roth et al. 1995). The *grk* transcripts and protein are positioned at the posterior pole of the oocyte in association with the germinal vesicle (Fig. 1.1.2.3A). The Grk signal is sent by the germinal vesicle to the posterior follicle cells. It is received by torpedo/DER in the follicle cells, a *Drosophila* homologue of the EGF receptor tyrosine kinase (Roth *et al,* 1995). The signal activates a signal transduction pathway in the follicle cells involving *Ras, D-raf, and Dsor1*. In response the follicle cells adopt a posterior fate and respond with an undefined signal that leads to the repolarisation of the oocyte microtubules (Fig. 1.1.2.3B). In the early stages of oogenesis the microtubule network is used to transport mRNAs and proteins to the previtellogenic oocyte. The reorientation of the microtubule network at stage 7 leads to relocalisation of the maternal determinants *bicoid* and *oskar* to the anterior and posterior pole respectively, a crucial movement for the establishing of the A/P axis(Fig. 1.1.2.3B-C). The repolarisation of the MTOC also leads to redistribution of grk mRNAs and protein and migration of the oocyte nucleus to the dorsal-anterior position, where later they activate a number of genes required for D/V patterning of the eggshell. Protein kinase A (PKA) is involved in the processes of disassembling and reassembling of MTOC (Lane and Kalderon 1995). The product of the gene *mago nashi* is also required for the microtubule reorganisation (Micklem et al. 1997). The

anterior determinant *bicoid (bcd)* encodes a transcription factor with a DNA binding site. It is localised to the anterior pole of the oocyte at stage 7 and is translated after fertilisation of the egg to produce an anterior to posterior gradient in the embryo (Berleth et al. 1988; St Johnston et al. 1989). It has been found that three genes are directly involved in *bcd* RNA positioning: *exuperantia* (Wang and Hazelrigg 1994), *swallow* (Stephenson et al. 1988), and *staufen* (St Johnston et al. 1991).

The differentiation of the follicle cells as posterior cells also requires signalling among the somatic cells. It has been found that the pre-differentiation of the polar follicle cells is important for this process, which indicates that pathways such as *Notch-Delta* and *Hh* are involved in axis determination (Forbes et al. 1993)

1.1.2.4 Dorsal/Ventral axis determination

A crucial movement for the establishment of the Dorsal/Ventral polarity is the migration of the germline vesicle to the anterior region in a microtubule dependent manner at stage 8 (Fig. 1.1.2.3C). As the germinal vesicle moves (Fig. 1.1.2.3D) it remains associated with the *grk* RNA and protein, which leads to their localisation at the dorsal-anterior part of the oocyte. That induces the follicle cells facing the germline vesicle to adopt a dorsal fate. The subsequent signalling is similar to that determining A/P axis, involving the top/DER, *Ras*, *D-raf*, and *Dsor1* pathways (Fig. 1.1.2.3E).

The gene *fs(1)k10* is required for the localisation of grk RNA to the dorsal-anterior region. It encodes a helix-loop-helix DNA binding protein (Schupbach 1987). The *k10* mRNA localises specifically to the oocyte nucleus requiring *capucino* and *spire* activity. Several other genes are needed for *grk* RNA localisation, such as: *orb* (Christerson and McKearin 1994; Lantz et al. 1994), *Bic-D* (Swan and Suter 1996), and *squid* (Norvell et al. 1999).

The transfer of the Grk signal requires the function of *cornichon (cni)* and *braniac (brn)* genes. Cornichon is a small 144 aa protein with a characteristic large hydrophobic domain, which suggests a role for *cni* in the localisation of the grk protein to the membrane (Roth et al. 1995). The neurogenic gene *brn* is required for the maintenance of the follicular epithelium in the established egg chamber, and for regulation of grk-top/DER interactions. Several genes have been found to be

expressed downstream of the grk-top/DER signalling pathway. A putative transmembrane protein rhomboid was the first gene identified to regulate the signalling pathway. CF is a transcription factor active in the ventral follicle cells only, it is required for the formation of the eggshell.

The *Broad Complex* is a gene responsible for the morphogenesis of the dorsal appendages and is found to be downstream of the grk-top/DER pathway. Br-C proteins, zinc-finger transcription factors, are expressed in a dynamic pattern in two groups of dorsal-anterior follicle cells at stage 13 of oogenesis. A gene *pointed (pnt),* that encodes a transcription factor with ETS domain, represses *Br-C* expression in the dorsal mid-line follicle cells.

In the absence of a dorsal signal (the follicle cells do not receive the torpedo mediated signal) the follicle cells adopt a ventral fate by default.

Fig. 1.1.2.3 Schematic representation of the main steps in the development of the AP and DV polarity during oogenesis (adapted from van Eeden and St Johnston 1999)

A: At stage 5-6 the oocyte contains an active MTOC (microtubule organising centre) which extends to the nurse cells and polarises the egg chamber (MTOC is labelled in Red). During stage 5-6 the oocyte nucleus and the GRK signal it produces are located at the posterior part of the oocyte. Grk signals to the adjacent follicle cells and induces them to adopt a posterior fate (Blue). At that time the follicle cell at the anterior tip of the oocyte adopts an anterior fate (Yellow).

B: During stages 6-7 the follicle cells send an undefined signal back to the oocyte. First this leads to disassembly of the MTOC and reorientation of the microtubules.

C: Then (stage 8) the oocyte nucleus move in a microtubule dependent manner to the anterior compartment of the oocyte.

D: At stage 9 *oskar* mRNA (Black) and Staufen are localised to the posterior of the oocyte while *bicoid* mRNA (Purple) act at the anterior tip of the oocyte. The microtubule cytoskeleton is required for this process.

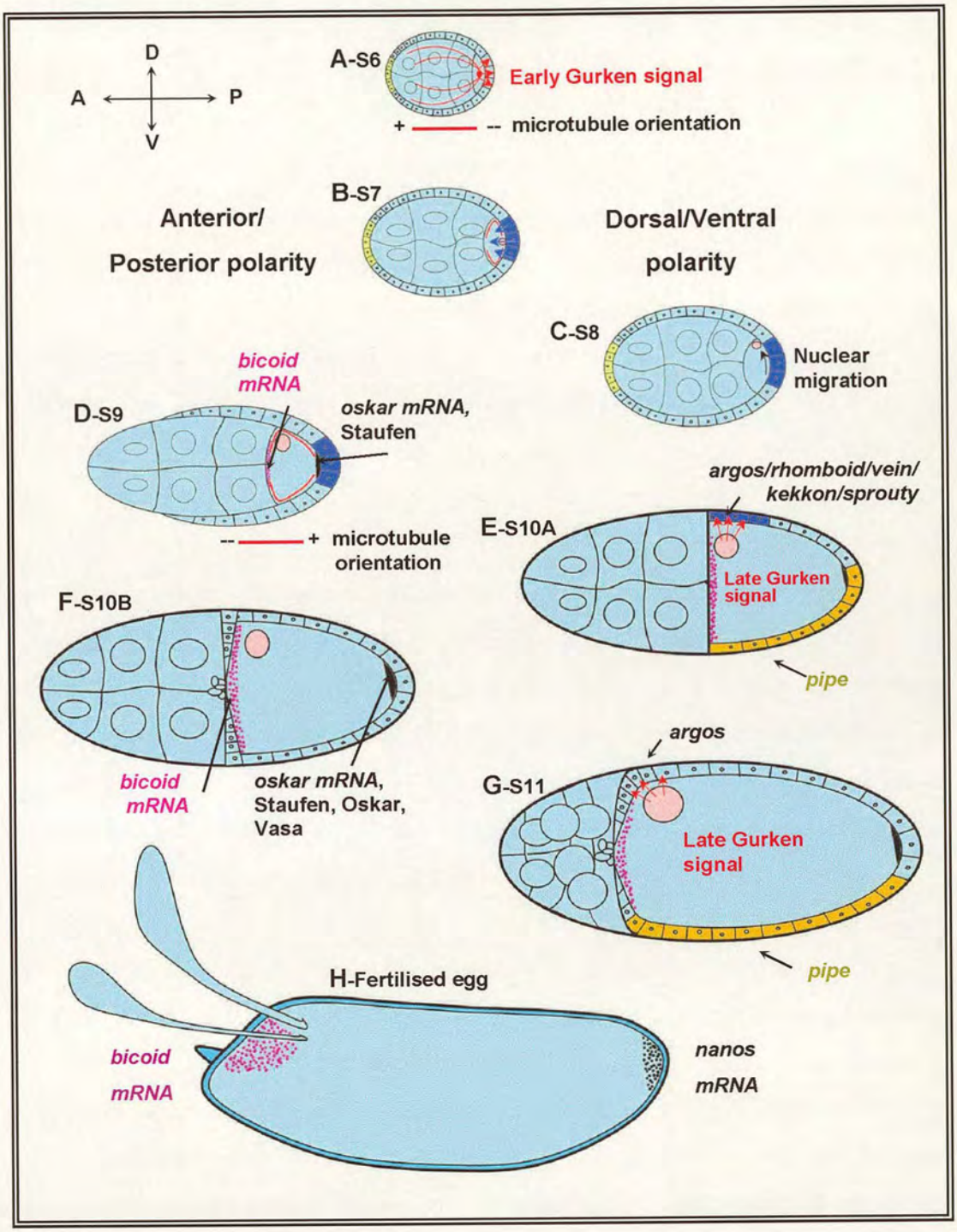
E: The late Grk signal activates the EGFR pathway in the adjacent follicle cells and induces a number of secondary genes, thus determining a dorsal fate. The activation of EGFR pathway restricts *pipe* expression (Orange) to the ventral follicle cells.

F: During stage 10B Oskar is required to anchor the *osk* mRNA to the posterior part of the oocyte.

G: The activation of EGFR by Spitz, Grk and Vein induces the dorsal expression of the EGFR inhibitor Argos. Argos represses the EGFR signal in the dorsal anterior follicle cells thus splitting it into two lateral peaks.

H: In the fertilised egg *bicoid* mRNA is actively translated into the cytoplasm to produce a gradient of the anterior morphogen. From stage 12 onwards *nanos mRNA* is also translated and specifies abdomen formation.

Fig. 1.1.2.3 Schematic representation of the main steps in the development of the AP and DV polarity during oogenesis



1.2 *Approaches used to isolate genes important for oogenesis*

1.2.1 Genetic and general molecular methods

Considerable work has been carried out to elucidate the complex mechanisms controlling the development of the oocyte. The standard approach to study *Drosophila* development involves identifying mutations that disrupt the process of interest. The target gene then is characterised both genetically and at the molecular level by cloning, sequencing, and expression analysis.

The best source for studying genes involved in oogenesis is female sterile mutations. These lead to abortive egg production or laying eggs which fail to develop normally. Several genetic screens have been done in *Drosophila* to isolate female sterile mutants (fs) (Swan et al. 2001; Lasko 1994; Schupbach and Wieschaus 1991; Orr et al. 1989; Komitopoulou et al. 1983).

A number of genes are difficult to identify using the standard method like those with a subtle function, with a high degree of redundancy or genes which have functions at more than one time during development. New techniques have been developed to overcome the difficulties. A widely used method to study important genes in oogenesis is generating heatshock inducible transgenic flies by P-element transformation. A heatshock treatment can rescue the lethal phenotype and can be a valuable tool when the investigated gene has multiple functions at different stages of oogenesis.

An alternative method is to produce mosaic egg chambers, either by FLP-mediated recombination (Theodosiou and Xu 1998; Fedorova et al. 2001) or by the dominant female sterile techniques (Perrimon and Gans 1983).

A new method for targeted gene knockout has been developed by Rong et al. (2002). A donor construct is created to carry DNA from the gene to be targeted. This is introduced into the fly genome by P-element transformation. Transgenic expression of a FLP (site specific recombinase) and I-SceI endonuclease are used to generate extrachromosomal linear DNA molecules *in vivo*. Subsequently these molecules recombine with the corresponding locus resulting in alterations of the fly genome.

1.2.2 Enhancer trap systems

In order to isolate genes involved in cell determination, inter and intracellular transport, cell-cell communications and temporal and spatial regulation during oogenesis, P-mediated enhancer detection has proved extremely useful (Bellen et al. 1990; O'Kane and Gehring 1987; Grossniklaus et al. 1989). The method uses a modified transposable element from a *Drosophila* P-element (Fig. 1.2.2A). This synthetic element carries the *Escherichia coli lac Z* gene placed under the control of a weak constitutive promoter. The transposase gene is removed from the P-element and replaced by a dominant marker gene *(white, w^+)*. A separate helper plasmid is required to provide a source of transposase. It is only present for a single generation, after which it is removed by genetic crossing. Flies with a transformed genome have red eyes and can be easily selected on white *(w+)* background.

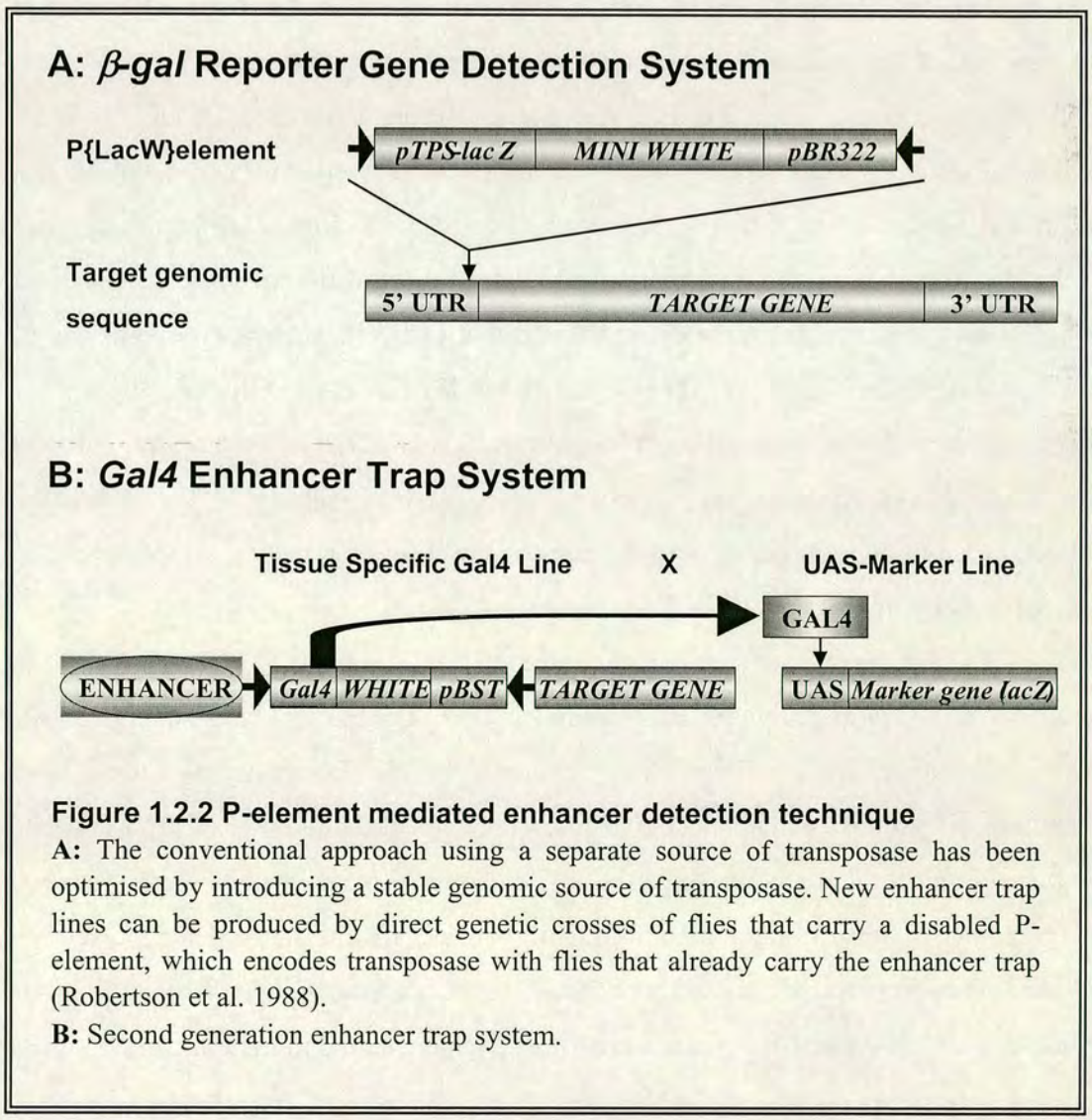


Figure 1.2.2 P-element mediated enhancer detection technique

A: The conventional approach using a separate source of transposase has been optimised by introducing a stable genomic source of transposase. New enhancer trap lines can be produced by direct genetic crosses of flies that carry a disabled P-element, which encodes transposase with flies that already carry the enhancer trap (Robertson et al. 1988).

B: Second generation enhancer trap system.

When the P-element is inserted near to an ovarian enhancer, it will be activated and β-gal is expressed in the same or a similar pattern to the adjacent gene that is normally regulated by that enhancer. Thus, *lac Z* acts as a tissue specific reporter gene. Expression patterns of interest can be identified and then the target gene cloned by excising the plasmid sequence of the P-element from the genome along with flanking DNA. It has been found that the P-elements exhibit some preference for inserting near the 5' end of genes, and they do not usually insert within heterochromatin. Berg and Spradling (1991); Kelley et al. (1987); Eggleston (1990) and Bownes (1990) showed that P-elements tend to insert into or near other P-elements, with a particular preference for bases 19-26 of the target P-element.

Second generation enhancer trap systems utilise yeast Gal-4 transcriptional activator in the P-element instead of the β-gal reporter gene (Brand and Perrimon 1993; Deng et al. 1997) (Fig.1.2.2B). The product of *Gal-4* can only activate expression of genes containing a Gal-4 binding site, known as UAS (upstream activation sequence). A second fly line contains a P-element construct with an UAS linked to a marker gene *(lac Z).* The progeny of a tissue specific Gal-4 line with a UAS-*lac Z* gene will express the marker in all cells in which the Gal-4 enhancer trap is active. This system can be a powerful tool for testing different marker genes in a single Gal-4 enhancer trap line. It was found that Gal-4 lines have no expression in the germline cells. That makes them a perfect marker for sets of follicle cells, e.g. for investigating cell-cell interactions in oogenesis. This system is far more potent than the original enhancer trapping system. It provides a possibility of Gal-4-directed misexpression and antisense expression to study the processes of follicle cell determination, migration, and differentiation.

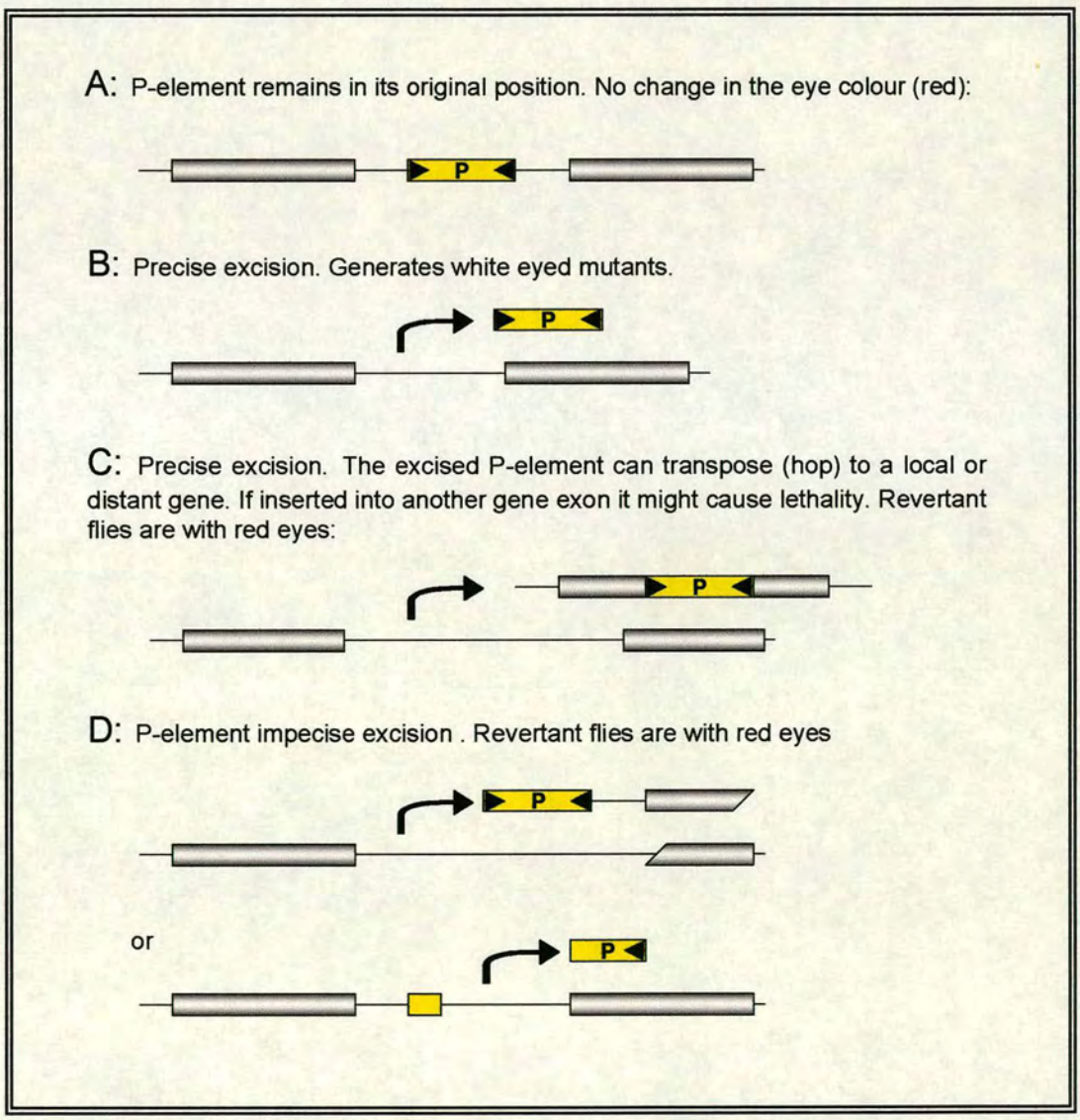
The P-element mediated enhancer trap system has proved extremely useful in studying *Drosophila* development. A wide variety of genes involved in the oogenesis have been identified by using this technique. In our laboratory we have been examining a number of enhancer trap lines for interesting patterns of reporter-gene expression during *Drosophila* oogenesis. These screens led to the isolation of several lines with dynamic staining patterns in the ovaries. In this thesis I am presenting studies on two unrelated genes: The *Broad Complex*, a gene isolated from a P[Gal-4] enhancer trap line and *Drosophila* Myosin V, a gene isolated in the course of P[lacZ]

screen, both of which proved to be crucial for normal oogenesis.

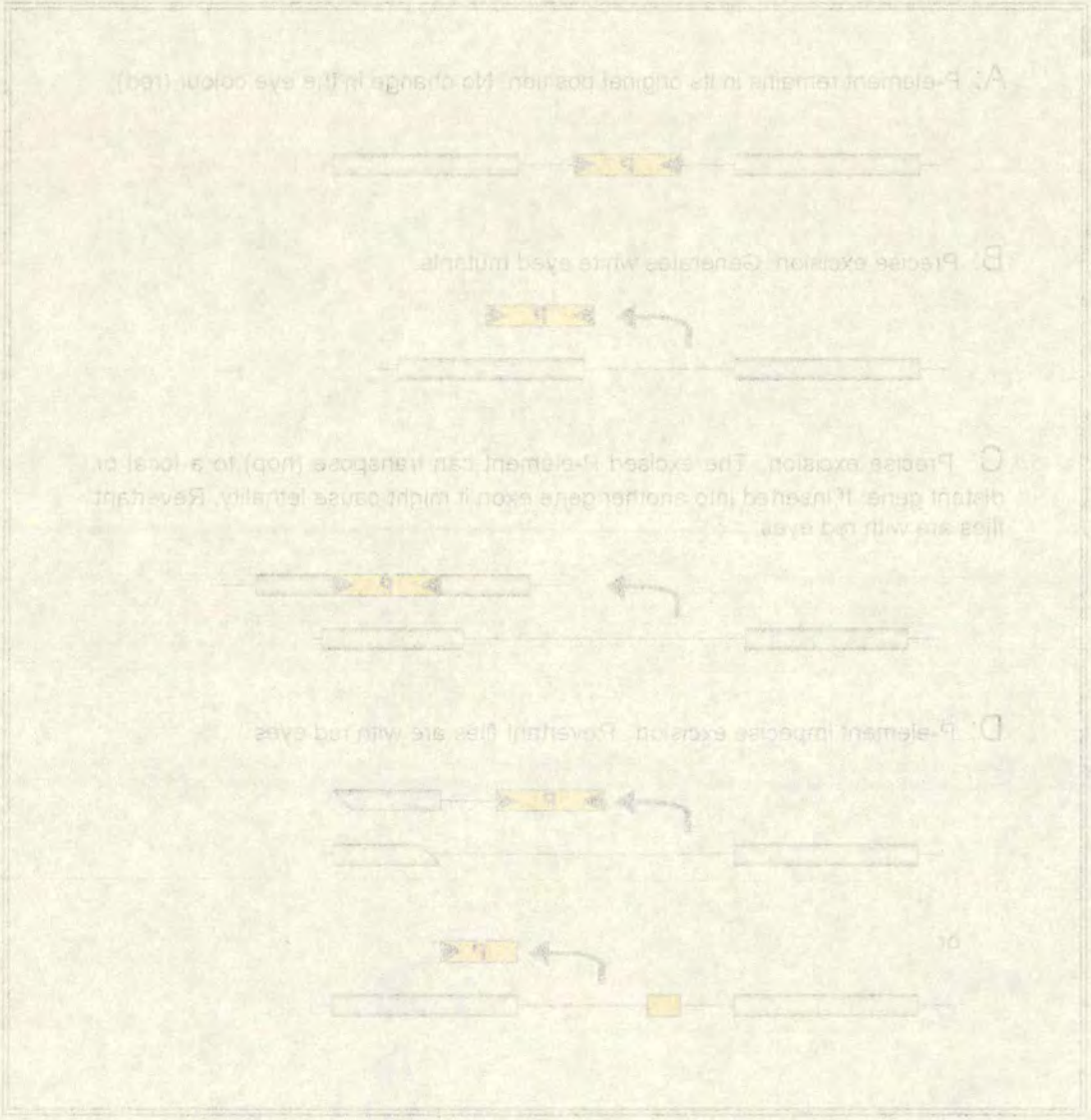
1.2.3 P-Hop mutagenesis

When a P-line of interest has a lethal phenotype it is crucial to show unequivocally that the P-element is responsible for the lethality. The P-element is mobilised by crossing with a fly strain that contains non-autonomous (non-mobile) P-element (expresses functional transposase). If the P-insertion causes the lethal phenotype, a precise excision should show a reversion to wild type (Fig. 1.2.3B)

Figure1.2.3 Possible outcomes from a mobilised P-element



A common problem in molecular genetics is obtaining mutations in genes known only by their DNA sequence. P-elements tend to insert into sites closely linked to the donor site (Golic 1994; Bownes 1990). If no P-element insertions are known for a given target gene a P-Hop mutagenesis can be undertaken. For this reason P-lines are selected with P- insertions close to the target gene or located within the noncoding sequences (introns) of the gene. When the P-element of the target line is mobilised by crossing with a flyline containing a non-mobile P-element, the target P-element can jump (hop) to a new position into a near by exon (Fig. 1.2.3C) or can delete a part of the target gene by imprecise excision, thus generating a mutation (Fig. 1.2.3D).



1.3 Project aims and outcomes

1.3.1 Molecular and functional analysis of *didum*, the *Drosophila* gene encoding unconventional Myosin V

As the oocyte develops it enlarges as a result of deposition of various products synthesised in the nurse cells and the fat body. These include proteins, mRNAs, and organelles essential for oogenesis and later embryonic development. A highly developed transport system is required for proper direction of these products to the target subcellular localisations. Genetic and molecular data has implicated myosin V in a variety of cell trafficking and cell movements. One of the aims of the current work is to understand the function of the Unconventional Myosin V in *Drosophila*. In this thesis its function in oocyte/egg production is analysed. To do this the distribution of Myosin V mRNA and protein in the cells and the genetic pathways they interact with to influence oogenesis are studied.

1.3.2 Identification, molecular and phylogenetic analysis of *Drosophila* myosin genes

Myosins constitute a large family of motor proteins that use the energy of ATP hydrolysis to move along actin filaments. Currently phylogenetic analysis places myosins into 18 classes based on specific features of their conserved head domain. How many myosins are needed to ensure the proper development and function of eukaryotic organism? Previously three types of myosins were found in budding yeast, six in the nematode *C. elegans*, and at least 12 in human. How many myosin genes does *Drosophila* have? Analysis of the completed *D. melanogaster* genome project have been performed to determine the number of myosin-encoding genes. The evolutionary relationships between members of the myosin family in *Drosophila* and other phylogenetically diverse species are examined. In addition we analysed the molecular structure of the newly identified myosin genes. Finally we analysed the domain composition of the proteins they encode and discuss their possible function.

1.3.3 Analysis of the function of *Broad Complex* during oogenesis

Over the last several years a number of developmentally important questions about cell activity and function of the somatic follicle cells have been studied. The data suggest that a crucial role is played by the follicle cells in oocyte determination, maturation and in polarity formation. However there are many more questions to be answered: how is the oocyte determined, what signals trigger follicle cell differentiation and what are the interactions between the subtypes of follicle cells and between the follicle cells and the germ line cells? The purpose of this work is to study the early function of *Br-C*, when it is expressed in all follicle cells at stage 6 of oogenesis. Its later expression in the dorsal-anterior follicle cells is related to its function in the formation of the chorion appendages (Deng and Bownes 1997). Since the chorion genes encode major eggshell components, and *Br-C* function has been reported to be necessary for chorion gene amplification during oogenesis (Orr et al. 1989; Huang and Orr 1992), we investigated the relationship between the *Br-C* and chorion gene amplification and expression.

Chapter Two: Materials and Methods

2.1 Materials

2.1.1 Chemicals and radioactive isotopes

General purpose chemicals were purchased from Sigma, BDH, Fisher Scientific and Fluka. A number of chemicals were supplied by other companies whose names will be stated accordingly in the text.

Isotopes, such as [^{32}P]dCTP, were obtained from Amersham Biosciences.

Digoxigenin labelling and detection reagents were purchased from Roche Molecular Biochemicals.

2.1.2 Restriction endonucleases and modifying enzymes

Restriction endonucleases were obtained from New England Biolabs and Promega.

Taq Polymerase and the accompanying buffers were supplied by Qiagen. High-fidelity Pfu Turbo DNA Polymerase was obtained from Stratagene. Klenow DNA polymerase was supplied by Roche Molecular Biochemicals.

T4 DNA Ligase, T4 Polynucleotide Kinase, CIP (Calf Intestinal Alkaline Phosphatase) and SAP (Shrimp Alkaline Phosphatase) were purchased from New England Biolabs.

T3, T7 and SP6 RNA Polymerases were obtained from Roche Molecular Biochemicals.

Reverse Transcriptase (SuperScript^{TM} II) was supplied by Invitrogen, Life Technologies.

2.1.3 Buffers and Solutions

All buffers and solutions were prepared with double distilled water (ddH_2O). Sterilisation, when required, was achieved by autoclaving for 30 minutes at 15 psi or by sterile filtering through a 0.22μm filter (Sartorius). All solutions were stored at room temperature (23-25°C) unless otherwise stated. General solutions and buffers are listed in Table 2.1.3. (A) , buffers for gel electrophoresis in Table 2.1.3 (B), staining reactives and detection reagents in Table 2.1.3 (C), and mounting media for microscopy in Table 2.1.3 (D).

Table 2.1.3 (A) List of general solutions and buffers

Solution	Components/Preparation	Final concentration
Borax	1.9g $Na_2B_4O_7.10H_2O$ (Mw: 381.4) Adjust the volume to 50ml with ddH_2O	100mM
Buffer FBBW	850µl 1M K-Phosphate Buffer, pH-6.8 1.5ml 2.5M KCl 250µl 5M NaCl 165µl 1M $MgCl_2$ 8ml Formaldehyde (Formalin) ~37% Adjust the volume to 50ml with ddH_2O. Sterilised through a 22µ filter.	17mM 75mM 25mM 3.3mM 6%
$CaCl_2$	73.5g $CaCl_2.2H_2O$ (Mw: 147.0) Adjust the volume to 100ml with ddH_2O	5M
Denhardt's (100x)	1g BSA (bovine serum albumin) 1g Ficoll (Wt: 400,000) 1g Polyvinylpyrolidone (Wt: 360,000) Adjust the volume to 50ml with ddH_2O	2% 2% 2%
DNA Hybridisation buffer	125ml 20x SSPE 25ml 10% SDS 25ml 100x Denhardt's 1ml 1mg/ml sssDNA (denatured) 325ml ddH_2O	5x 0.5% 5x 20µg/ml
DNA Hybrix (for *in-situ*)	25ml Deionised formamide (Mw: 45.04) 12.5ml 20x SSC 250µl dsssDNA [10mg/ml] 50µl [50mg/ml] Heparin 500µl 10% Tween-20 11.5ml ddH_2O	50% 5x 50µg/ml 50µg/ml 0.1%
dNTP mix (10mM each dNTP)	10µl 100mM dATP 10µl 100mM dCTP 10µl 100mM dGTP 10µl 100mM dTTP 60µl ddH_2O	10mM 10mM 10mM 10mM
EBR (Ephruzzi Beadle Ringer)	7.5g NaCl (Mw: 58.44) 0.35gKCl (Mw: 74.55) Adjust the volume to 1L with ddH_2O	0.75% 0.035%

Solution	Components/Preparation	Final concentration
EDTA, pH-8.0	93g $Na_2EDTA.2H_2O$ (Mw: 372.24) Adjust the pH to 8.0 with ~20g NaOH Adjust the volume to 500ml with ddH_2O	0.5M
EGTA, pH-8.0	1.9g EGTA (Mw: 380.4) Adjust the pH to 8.0 with 10N NaOH Adjust the volume to 10ml with ddH_2O	0.5M
FBS	2g Paraformaldehyde 5ml 10xBPS 45ml ddH_2O Dissolve at 65°C. Sterilise through a 22µ filter	4% 1x
Glycine	100mg Glycine (Mw: 75.07) in 1ml ddH_2O	100mg/ml
Glycine, pH-2.5	3.75g Glycine (Mw: 75.07) Adjust the pH to 2.5 with cHCl Adjust the volume to 500ml with ddH_2O	0.1M
HCl	8.6ml concentrated HCl Adjust the volume to 50ml with ddH_2O	2N
Heparin	50mg Heparin 800µl ddH_2O 200µl 20xSSC	50mg/ml 4x
K Acetate	78.59g CH_3COOK (Mw: 98.14) Adjust the volume to 100ml with ddH_2O	8M
K ferricianide	1.65g $K_3[Fe(CN)_6]$ (Wt: 329.25) (red) 50ml ddH_2O	0.1M
K ferrocyanide	2.11g $K_4[Fe(CN)_6].3H_2O$ (Wt: 422.4) (yellow) 50ml ddH_2O	0.1M
K_2HPO_4	17.42g K_2HPO_4 (Mw: 174.18) in 100ml ddH_2O	1M
KCl	7.46g KCl (Mw: 74.55) in 100ml ddH_2O	1M
KCl	18.64g KCl (Mw: 74.55) in 100ml ddH_2O	2.5M
KH_2PO_4	13.61g KH_2PO_4 (Mw: 136.09) in 100ml ddH_2O	1M
KOH	28.06g KOH (Mw: 56.11) in 100ml ddH_2O	5M
Levamisole	0.24g Levamisole (Mw: 240.8) 500µl H_2O 500µl Glycerol Adjust the volume to 1ml with ddH_2O	1M 50%

Solution	Components/Preparation	Final concentration
Lysozyme	50mg Lysozyme 1ml ddH_2O	50mg/ml
$MgCl_2$	9.5g $MgCl_2$ (Mw: 95.21) Adjust the volume to 100ml with ddH_2O	1M
$MgSO_4$	24.6g $MgSO_4.7H_2O$ (Mw: 246.5) Adjust the volume to 100ml with ddH_2O	1M
$MnCl_2$	39.58g $MnCl_2.4H_2O$ (Mw: 197.9) Adjust the volume to 100ml with ddH_2O	2M
MOPS, pH-6.8	20.9g MOPS (Mw: 209.3) Adjust the pH to 6.8 with dry NaOH. Adjust the volume to 100 ml with ddH_2O. Sterilise by filtration through a 0.22 µm filter	1M
Na Acetate, pH-4.0	13.6g $CH_3COONa.3H_2O$ (Mw: 136.08) Adjust the pH to 4 with Glacial Acetic acid Adjust the volume to 50ml with ddH_2O	2M
Na Acetate, pH-5.2	40.8g $CH_3COONa.3H_2O$ (Mw: 136.08) Adjust the pH to 5.2 with Glacial Acetic acid Adjust the volume to 100ml with ddH_2O	3M Na 5M Acetate
Na Citrate, pH-7.5	14.7g $Na_3C_6H_5O_7.2H_2O$ (Mw: 294.1) Adjust the pH to 7.5 with Citric acid Adjust the volume to 50ml with ddH_2O	1M
NaCl	146.1g NaCl (Mw: 58.44) in 500ml ddH_2O	5M
NaOH	80g NaOH (Wt: 40.0) Adjust the volume to 20ml with ddH_2O	10N
NaOH	16g NaOH (Wt: 40.0) Adjust the volume to 1L with ddH_2O	0.4N
PBS (10x), pH-7.5	40g NaCl (Mw. 58.44) 5.7g Na_2HPO_4 (Mw. 142.0) 1.2g NaH_2PO_4 (Mw: 120) Adjust the pH to 7.4 with 10N NaOH Adjust the volume to 500ml with ddH_2O	1.37M 80mM 20mM
PBT (1x)	50ml 10xPBS 450ml ddH_2O 500µl Tween 20	1x 0.1%

Solution	Components/Preparation	Final concentration
Phenol/Chloro-form*	50ml Phenol (saturated with TE Buffer, pH-8.0) 50ml Chloroform	
Phosphate buffer, pH-6.8	8.66g K_2HPO_4 (Mw: 174.18) 6.8g KH_2PO_4 (Mw: 136.09) in 100ml ddH_2O	1M
PMS	2g Paraformaldehyde 40ml H_2O 30µl 10M NaOH Dissolve at 65°C. Then add: 5ml 0.5M PIPES, pH-7.0 100µl 0.5M EGTA, pH-8.0 Adjust pH to 6.8 with 1N HCl, then add: 100µl 1M $MgSO_4$ Adjust the volume to 50ml with ddH_2O	4% 50mM 1mM 2mM
Proteinase K	20mg Proteinase K 10µl 1M Tris-HCl, pH-7.5 590µl ddH_2O 400µl Glycerol	20mg/ml 10mM
RbCl	12.09g RbCl (Mw: 120.9) in 100ml ddH_2O	1M
Ringer's saline for *Drosophila*, pH-7.2	6.8g KCl (Mw: 74.55) 1.35g NaCl (Mw: 58.44) 0.17g $CaCl_2.2H_2O$ (Mw: 147) 0.6g Tris-base (Mw: 121.14) Adjust the pH to 7.2 with 1N HCl. Adjust the volume to 500ml with ddH_2O. Sterilise through a 0.22µm filter.	180mM 45mM 2.3mM 10mM
RNA Hybridisation buffer	125ml Formamide (Mw: 45.04) 62.5ml 20xSSC 5ml 1M Na-Phosphate buffer, pH-6.8 25ml 10% SDS 12.5ml 100x Denhardt's 2mg total yeast RNA or sssDNA (denatured) Adjust the volume to 500ml with ddH_2O	50% 5x 20mM 1% 5x 8µg/ml

Solution	Components/Preparation	Final concentration
RNA Hybrix, pH-6.5 (for *in-situ*)	25ml Deionized formamide (Mw: 45.04) 12.5ml 20x SSC 100µl [50mg/ml] tRNA (RNase free) 50µl [50mg/ml] Heparin 500µl 10% Tween-20 12ml ddH_2O Adjust the pH to 6.5 with HCl	50% 5x 100µg/ml 50µg/ml 0.1%
RNase (DNase free)	Dissolve 10mg RNase in 1ml. Aliquot into 4 eppendorfs. Heat up to 110°C for 20min at the heating block. Leave in the bloke to cool down gradually to RT. Store at –20°C.	10mg/1ml
SBE buffer	20g $NaHCO_3$ (Mw: 84.01) 2ml 0.5M EDTA, pH-8.0 Adjust the volume to 1L with ddH_2O	2% 1mM
SDS	50g SDS (Mw: 288.38) Adjust the volume to 500ml with ddH_2O	10%
SSC (20X), pH-7.0	87.65g NaCl (Wt: 58.44) 44.1 $Na_3C_6H_5O_7.2H_2O$ (Wt: 294.1) Adjust the pH to 7.0 with 10N NaOH Adjust the volume to 500ml with ddH_2O	3M 300mM
SSPE, pH-7.4	87.65g NaCl (Mw: 58.44) 15.6g NaH_2PO_4 (Mw: 156.01) 4.7g $Na_2EDTA.2H_2O$ (Mw: 372.24) Adjust the pH to 7.4 with ~ 3ml 10N NaOH Adjust the volume to 500ml with ddH_2O	3M 200mM 25mM
sssDNA	10mg Sonicated salmon sperm DNA 1ml ddH_2O	10mg/ml
Sucrose	8.56g Sucrose (Mw: 342.3) in 100ml ddH_2O	250mM
TB Solution	0.3g PIPES (Mw: 302.4) 2.75ml 2M $MnCl_2$ 300µl 5M $CaCl_2$ 10ml 2.5M KCl Adjust the pH to 6.7 with 10N KOH prior to adding the $MnCl_2$!!! Adjust the volume to 100ml with ddH_2O	10mM 55mM 15mM 250mM

Solution	Components/Preparation	Final concentration
TBS (10x)	12.1g Tris Base (Mw: 121.14) 40g NaCl (Mw: 58.44) Adjust the pH to 7.6 with cHCl Adjust the volume to 500ml with ddH_2O	200mM 1.37M
TBS-T (1x)	50ml TBS (10x) 450ml ddH_2O 500µl Tween-20	1x 0.1%
TE, pH-8.0	1ml 1M Tris, pH-8.0 200µl 0.5M EDTA, pH-8.0 in 100ml ddH_2O	10mM 1mM
Tris-HCl, pH-7.4 pH-7.6 pH-8.0	60.55g Tris Base (Mw: 121.14) For pH-7.4, add ~35ml cHCl For pH-7.6, add ~30ml cHCl For pH-8.0, add ~20ml cHCl Adjust the volume to 500ml with ddH_2O	1M
Triton X-100	0.5ml Triton X-100 Adjust the volume to 500ml with ddH_2O	0.1%
Triton X-100	5ml Triton X-100 Adjust the volume to 50ml with ddH_2O	10%
tRNA (RNase free)	50mg tRNA in 1ml ddH_2O	50mg/ml
Tween 20	5ml Tween 20 Adjust the volume to 50ml with ddH_2O	10%

* This Phenol/Chloroform mix is for DNA deproteinisation. The phenol for RNA purification is acidic, with pH-4.3.

Table 2.1.3 (B) List of loading buffers and running buffers for gel electrophoresis

Buffer	Components/Preparation	Final concentration
DNA Sample buffer (6x)	3ml 0.1% Bromphenol Blue	0.03%
	3ml Glycerol	30%
	4ml ddH_2O	
MOPS buffer (10x), pH-7.0	20.9g MOPS (Mw: 209.3)	200mM
	5.44g CH_3COONa (Mw: 136.08)	80mM
	10ml 0.5M EDTA, pH-8.0	10mM
	Adjust the pH to 7.0 with 2N NaOH.	
	Adjust the volume to 500ml with ddH_2O	
	Sterilise through a 0.22μm filter.	
PIPES, pH-6.5 pH-7.0	15.1g PIPES (Mw: 302.4)	1M
	Adjust the pH to 6.5 or 7.0 as needed with cHCl	
	Adjust the volume to 50ml with ddH_2O	
Protein Sample buffer (2x)	1.25ml 1M Tris-HCl, pH 6.8	125mM
	2ml Glycerol	10%
	4ml 10% SDS	4%
	1ml 14.4M β-Mercaptoethanol (Mw: 78.01)	1.44M
	1ml 0.1% Bromphenol Blue	0.01%
	0.75ml ddH_2O	
RNA Sample buffer (1.5x)	500μl Deionised formamide (Mw: 45.04)	50%
	100μl 10x MOPS	1x
	270μl 37% Formaldehyde	10%
	100μl Glycerol	10%
	30μl 0.1% Bromphenol Blue	0.003%
	2μl 10mg/ml EtBr	20μg/ml
TAE (50x)	121g Tris Base (Mw: 121.14)	2M
	28.55ml Glacial Acetic acid	
	50ml 0.5M EDTA, pH-8.0	50mM
	Adjust the volume to 500ml with ddH_2O	
TGS buffer (10x) (Running buffer for PAGE)	15.15g Tris base (Mw: 121.1)	250mM
	72g Glycine (Mw: 75.07)	1.92M
	5g SDS (Mw: 288.38)	1%
	Adjust the volume to 500ml with ddH_2O	

Table 2.1.3 (C) List of colour reactives and detection reagents

Name	Components/Preparation	Final concentration
Coomassie gel destain	100ml Methanol 100ml Glacial Acetic acid 800ml ddH_2O	10% 10%
Coomassie gel stain	1g Coomassie Blue R-250 450ml Methanol 450ml ddH_2O 100ml Glacial Acetic acid	0.1% 45% 10%
DAB (10x)	20mg DAB-4HCl (Diaminobenzydine tetrahydochloride) 10ml 1x PBS	2mg/ml
DAB Staining solution	900µl 1x PBS 100µl 10x DAB solution 1µl 30% H_2O_2	1xPBS 0.2mg/ml 0.03%
Galactosidase staining buffer	500µl 1M Phosphate buffer, pH-7.0 3ml 0.1M $K_3[Fe(CN)_6]$ (or 99mg) 3ml 0.1M $K_4[Fe(CN)_6]$ (or 127mg) 1.5ml 5M NaCl 50µl 1M $MgCl_2$ Adjust the volume to 50ml with ddH_2O	10mM 8mM 8mM 150mM 1mM
Hoechst (1000x)	1mg Hoechst 33258 1ml 1x PBS	1mg/ml
IPTG	120mg Isopropyl β-D-thiogalactoside (Mw: 238.3) Adjust the volume to 1ml with ddH_2O	0.5M
Maleic buffer, pH-7.5	5.8g Maleic Acid (Mw: 116.1) 4.3g NaCl (Mw: 58.44) Adjust the pH to 7.5 with dry NaOH Adjust the volume to 500ml with ddH_2O	100mM 150mM
NBT	75mg 4-Nitro blue tetrazolium chloride 700µl N,N'-Dimethylformamide (NBT) (Mw: 73.09) 300µl ddH_2O	75mg/ml 70%
Phalloidin, rhodamine conjugated (20x)	100µg Phalloidin, rhodamine conjugated 400µl 99.7% EtOH	200µg/ml =150µM
Name	**Components/Preparation**	**Final**

		concentration
Ponceau S (10x)	2g Ponceau S 30g Trichloracetic acid (Mw: 163.39) 30g 5-Sulfosalicylic acid dihydrate (Mw: 254.22) Adjust the volume to 100ml with ddH_2O	2% 30% 30%
Staining solution	5ml 1M Tris-HCl, pH-9.5 1ml 5M NaCl 2.5ml1M $MgCl_2$ 500µl 10% Tween 20 41ml ddH_2O	100mM 10mM 50mM 0.1%
p-Coumaric acid Stock	0.15g p-Coumaric acid (Mw: 164.2) 10ml DMSO	90mM
Luminol Stock	0.44g Luminol (Mw: 177.2) 10ml DMSO	250mM
ECL Reagent	250µl Luminol Stock 110µl p-Coumaric acid Stock 5ml 1M Tris-HCl, pH- 8.5 Adjust the volume to 50ml with ddH_2O	1.25mM 0.2mM 100mM
X-Gal	50mg X-Gal (5-Bromo-4-chloro-3-indolyl β-D-galactoside, Wt: 408.64) 1ml N,N'-Dimethylformamide (Mw: 73.09)	50mg/ml =0.3M
X-Gal	80mg X-Gal (5-Bromo-4-chloro-3-indolyl β-D-galactoside, Wt: 408.64) 1ml N,N'-Dimethylformamide (Mw: 73.09)	8%
X-Phosphate (BCIP)	50mg X-Phosphate (5-Bromo-4-chloro-3-indolyl phosphate p-toluidine salt 1ml N,N'-Dimethylformamide (Mw: 73.09)	50mg/ml

Table 2.1.3 (D) List of mounting media for conventional and fluorescent microscopy

Media	Components/Preparation	Final concentration
FISH media	5ml 1M Tris-HCl, pH-8.5 1.25g DABCO (Mw: 112.2) 45ml Glycerol	100mM 2.5% 90%
Glycerol/PBS, pH-7.4	5ml 10x PBS, pH-7.4 40ml Glycerol 5ml ddH_2O	1x 80%
Glycerol/PIPES, pH-6.5	2.5ml 1M Pipes, pH-6.5 40ml Glycerol 7.5ml ddH_2O	50mM 80%
Hoyer's media	6g Arabic gum in 10ml ddH_2O Dissolve completely O/N. Then add: 4g Glycerol 40g Chloral hydrate (Mw: 165.4) Centrifuge for 2 hours at maximum speed and collect the cleared supernatant	15% 10% 100%
Hoyer's mount media	20ml Hoyer's media 20ml Concentrated Lactic acid	50% 50%

2.1.4 Media preparation

All types of media (Table 2.1.4) were prepared by Institute media staff and sterilised by autoclaving (15 minutes at 15 psi). Subsequently the media was stored at room temperature.

Table 2.1.4 Media list

Media type	Components / Preparation	Final concentration
LB (Luria Broth)	10g NaCl 10g Bacto-tryptone 5g Yeast extract Adjust the volume to 1L with ddH_2O	1% 1% 0.5%

Media type	Components / Preparation	Final concentration
LB agar	10g NaCl	1%
	10g Bacto-tryptone	1%
	5g Yeast extract	0.5%
	15g Difco agar	1.5%
	Adjust the volume to 1L with ddH_2O	
NZY Top Agar	1L of NZY+	
	7g Agarose**	0.7%
NZY Agar	1L of NZY+	
	15g Agar	1.5%
NZY^+ Broth*	5g NaCl	0.5%
	2g $MgSO_4.7H_2O$ (Mw: 246.5)	8mM
	10g NZ amine (casein hydrolysate)	1%
	5g Yeast extract	0.5%
	Adjust the pH to 7.5 with 5N NaOH.	
	Adjust the volume to 1L with ddH_2O	
SM Buffer (Phage buffer)	5.8g NaCl (Mw: 58.44)	100mM
	2.0g $MgSO_4.7H_2O$ (Mw: 246.5)	8mM
	50ml 1M Tris-HCl, pH-7.5	50mM
	5ml 2% (w/v) gelatine	0.01%
	Adjust the volume to 1L with ddH_2O	
SOB^+ Broth*	20g Bacto tryptone	2%
	5g Yeast extract	0.5%
	0.58g NaCl (Mw: 58.44)	10mM
	1ml 2.5M KCl	2.5mM
	5g $MgSO_4.7H_2O$ (Mw: 246.5)	20mM
	Adjust pH to 7.0 with 5N NaOH	
	Adjust the volume to 1L ddH_2O	
SOC^+ Broth*	1 L of SOB Broth	
	Autoclave. Add:	
	10ml 2M filter sterilised Glucose	20mM

* "+" Indicates that Mg salt is added to the media.
**Agarose is used in place of agar since agar can inhibit enzyme activity. This might present a problem when analysing the phage colonies later.
***In event of mite infection, strips of Whatman 3MM filter paper soaked in 3% (v/v) Benzyl Benzoate in Ethanol and air dried are placed on top of the food.

2.1.5 Antibiotics

Stock solutions from all antibiotics were prepared at the recommended concentrations (Table 2.1.5) and stored at -20°C.

Table 2.1.5 Antibiotics list

Antibiotic	Stock solution concentration	Working concentration
Ampicillin	50mg/ml	50μg/ml
Chloramphenicol	68mg/ml	12.5μg/ml* or 50μg/ml
Kanamycin	25mg/ml	50μg/ml
Tetracycline	5mg/ml	50μg/ml

* Chloramphenicol in lower working concentration is required for selection and growing of ESTs strains.

2.2 Bacterial strains and plasmids

Table 2.2 (A) List of bacterial strains

Host strain	Genotype/Source
BL21(DE3) pLysS (Stratagene)	F *omp*T *hsd*S_B(r_B m_B) *gal dcm* λ (DE3) [pLysS Cam^r]
SCS110 (Stratagene)	*rps*L (Str^r) *thr leu end*A *thi*-1 *lac*Y *gal*K *gal*T *ara ton*A *tsx dam dcm sup*E44 Δ*(lac-pro*AB*)* [F′ *tra*D36 *pro*AB *lac*I^{q}ZM15]
SOLR (Stratagene)	e14 (mcrA) Δ*(mcrCB-hsd*SMR-*mrr)*171 *sbc*C *rec*B *rec*J *uvr*C *umu*C::Tn*5* (Kan^r) *lac gyr*A96 *rel*A1 *thi*-1 *end*A1 λ^R [F' *pro*AB *lac*I^{q}ZΔM15*] Su*$^-$ (nonsuppressing)
Top 10 (Invitrogen)	*F mcr*A Δ*(mrr-hsd*RMS-*mcr*BC*)* Φ80 *lac*ZΔM15 Δ*lac*X74 *deo*R *rec*A1 *ara*D1*39* Δ *(ara-leu)*7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G
XL 1 Blue (Stratagene)	*rec*A1 *end*A1 *gyr*A96 *thi*-1 *hsd*R17 *sup*E44 *rel*A1 *lac* [F′ *pro*AB *lac*I^{q}ZΔM15 Tn*10* (Tet^r)]
XL 10-Gold (Stratagene)	Tet^r Δ*(mcr*A*)*183 Δ*(mcr*CB-*hsd*SMR-*mrr)*173 *end*A1 *sup*E44 *thi*-1 *rec*A1 *gyr*A96 *rel*A1 *lac* Hte [F' *pro*AB *lac*I^{q}ZΔM15 Tn*10* (Tet^r) Amy Cam^r]
XL1-Blue MRF' (Stratagene)	Δ*(mcr*A*)*183 Δ*(mcr*CB-*hsd*SMR-*mrr)*173 *end*A1 *sup*E44 *thi*-1 *rec*A1 *gyr*A96 *rel*A1 *lac* [F' *pro*AB *lac*I^{q}ZΔM15 Tn*10* (Tet^r)]

Table 2.2 (B)

Construct name	Insert	Vector	Cloning enzymes	Selec-tion	Strain
pM1	Myosin V cDNA clone	pBluescript SK- (Stratagene)	EcoRI+XhoI	Amp^R	XL1 Blue
pM2	Myosin V cDNA clone	pBluescript SK- (Stratagene)	EcoRI+XhoI	Amp^R	XL1 Blue
pM3	Myosin V cDNA clone	pBluescript SK- (Stratagene)	EcoRI+XhoI	Amp^R	XL1 Blue
pM4	Myosin V cDNA clone	pBluescript SK- (Stratagene)	EcoRI+XhoI	Amp^R	XL1 Blue
pSK-	n/a	pBluescript SK- (Stratagene)	n/a	Amp^R	XL1 Blue
pGSK 7.6 kb	pBluescript SK- (Stratagene)	pEGFP-N1 (Clontech)	KpnI+HindIII	Amp^R + Kan^R	XL1 Blue
pRSK 7.6 kb	pBluescript SK- (Stratagene)	pEGFP-R1 (Clontech)	KpnI+HindIII	Amp^R + Kan^R	XL1 Blue
pGFT*	Myosin V Tail (sense)	pGSK	SpeI+AgeI	Kan^R	SCS110
pRFW*	Mayosin V ORF (sense)	pGSK	SpeI+AgeI	Kan^R	SCS110
TOPO	n/a	pCR II-TOPO (Invitrogen)	n/a	Kan^R	Top 10
7B7	c15 Chorion Protein (Spradling et al. 1980)	pBR322 (New England Biolabs)	EcoRI	Amp^R	XL1 Blue
7C8	c38 Chorion Protein (Spradling et al. 1980)	pBR322 (New England Biolabs)	EcoRI	Amp^R	XL1 Blue
Br-C Z1	Broad Complex Isoform Z1 (Bayer et al. 1996)	pCR1000 (Invitrogen)	TA cloning	Amp^R	XL1 Blue
Br-C Z2	Broad Complex Isoform Z2 (Bayer et al. 1996)	pCR1000 (Invitrogen)	TA cloning	Amp^R	XL1 Blue
Br-C Z3	Broad Complex Isoform Z3 (Bayer et al. 1996)	pCR1000 (Invitrogen)	TA cloning	Amp^R	XL1 Blue

Br-C Z4	Broad Complex Isoform Z4 (Bayer et al. 1996)	pCR1000 (Invitrogen)	TA cloning	Amp^R	XL1 Blue
Bs-cD5	Broad Complex Core domain (Bayer et al. 1996)	pBluescript SK- (Stratagene)	EcoRI	Amp^R	XL1 Blue
p32*	Myosin V Tail (anti-sense) (MacIver et al. 1998)	pRSET-C (Invitrogen)	EcoRI+DraI fragment in PvuII site	Amp^R	BL21DE3 pLysS
p33*	Myosin V Tail (sense) (MacIver et al. 1998)	pRSET-C (Invitrogen)	EcoRI+DraI fragment in PvuII site	Amp^R	BL21DE3 pLysS
SetC	n/a	pRSET-C (Invitrogen)	n/a	Amp^R	BL21DE3 pLysS
pUGT	Myosin V Tail–GFP (pGFT)	pUASp (Rorth 1998)	SpeI+XbaI	Amp^R	Top 10
pUGW	Myosin V ORF–GFP (pGFW)	pUASp (Rorth 1998)	SpeI+XbaI	Amp^R	Top 10
pUN2	pEGFP-N1 (Clontech)	pUASp (Rorth 1998)	KpnI+XbaI	Amp^R	XL1 Blue
pUR3	pEGFP-R1 (Clontech)	pUASp (Rorth 1998)	KpnI+XbaI	Amp^R	XL1 Blue
N110	n/a	pEGFP-N1 (Clontech)	n/a	Kan^R + Neo^R	SCS110
R110	n/a	pEGFP-R1 (Clontech)	n/a	Kan^R + Neo^R	SCS110
GFP-N1	n/a	pEGFP-N1 (Clontech)	n/a	Kan^R + Neo^R	XL1 Blue
GFP-R1	n/a	pEGFP-R1 (Clontech)	n/a	Kan^R + Neo^R	XL1 Blue
pUASp	n/a	pUASp (Rorth 1998)	n/a	Amp^R	XL1 Blue
pMSH	Myosin V (mutant fragment)	TOPO-3519 bp (Invitrogen)	Blunt end cloning	Kan^R	Top 10

pWTSH	Myosin V (wild type fragment)	TOPO-3519 bp (Invitrogen)	Blunt end cloning	Kan^R	Top 10
pHSMR	Myosin V ORF (antisense)	pCaSpeR-hs (Thummel and Pirrotta 1992)	XbaI	Amp^R	XL 10-Gold
pHSMS	Myosin V ORF (sense)	pCaSpeR-hs (Thummel and Pirrotta 1992)	XbaI	Amp^R	XL 10-Gold
pUAM	Myosin V ORF (sense)	pUASp (Rorth 1998)	SpeI+HpaI	Amp^R	XL 10-Gold
HS-Casper	n/a	pCaSpeR-hs (Thummel and Pirrotta 1992)	n/a	Amp^R	XL1 Blue

* Fusion constructs. In brackets is indicated the orientation of the subcloned gene, either sense or antisense.

2.2.1 Growing bacteria on agar plates

Bacterial cells are streaked onto agar plates to obtain an independent isolate. This is done to reduce the likelihood of working with a culture which has become contaminated and/or has accumulated mutations. The cells can be streaked from another plate, a stab or from a frozen glycerol stock.

Bacteria were picked by a sterile inoculating loop and streaked onto a LB plate in a sequential manner. A flamed glass spreader was used when larger volumes of cells (e.g. transformed competent cells) were to be plated. Subsequently plates were grown O/N at 37°C.

2.2.2 Growing bacteria in liquid cultures

Routinely, small scale bacterial cultures were required. Overnight cultures were prepared by inoculating 5ml of LB with a single bacterial colony. The appropriate antibiotic was added and the cultures grown O/N at 37°C with shaking (150rpm).

Larger scale cultures were prepared by subculturing an O/N culture in 50-500ml fresh LB. The appropriate antibiotic was added and the cultures grown at 37°C with shaking (200-250rpm) until the required OD_{590} was attained.

2.2.3 Storage of bacterial cultures

For short term storage, single colony was streaked on a LB plate supplemented with the appropriate antibiotic (Table 2.1.5). The culture was grown overnight at 37°C and stored at 4°C for up to one month.

For long term storage (up to five years) bacterial stabs were prepared. Toothpicks were used to pick and stab a single colony into a vial with LB-broth supplemented with 0.6% agar. The stabs were incubated for 6 hours at 37°C and stored at RT in dark.

Glycerol stocks were prepared for indefinite storage of the bacterial strains. 800μl of an O/N culture was added to 200μl sterile glycerol in a sterile 2ml screw-top tube. The mixture was homogenised by vortexing and stored at -80°C.

2.3 Manipulation of bacteriophage

2.3.1 Preparation of plating cells. Plating of bacteriophage library

A small amount of the bacterial glycerol stock, XL1 Blue MRF' was streaked onto agar plates. The plates were incubated O/N at 37°C. A single colony was picked and grown in 50ml LB supplemented with 0.5ml 20% Maltose (0.2% final concentration) and 0.5ml 1M $MgSO_4$ (10mM final concentration). The culture was incubated O/N at 30°C with shaking (the lower temperature keeps the bacterial cells from overgrowing). Alternatively the culture was grown for 5-6 hours (during the day) at 37°C with shaking. The culture was monitored spectrophotometrically to ensure the cells did not grow past OD_{600} of 1.0. The cells were pelleted down by centrifugation (500x g, 10 minutes). The supernatant was decanted and the cells gently resuspended in ~10ml 10mM $MgSO_4$ to give an OD_{600} of 0.5. The cells were stored at 4°C for up to 1 week.

2.3.2 Determination of the bacteriophage library titer.

A series of phage dilutions ($1:1x10^4$, $1:1x10^5$, $1:1x10^6$, $1:1x10^7$) were made in SM buffer. 1µl of each dilution was used to inoculate 200µl of Plating cells. After a 15 minute incubation, 3ml of NZY top agar (preheated to 48°C) were added and the mix plated onto NZY plates (preheated to 37°C). The plates were incubate O/N at 37°C. On the following day the number of plaques was counted and the plaque forming unit (pfu)/ml concentration was determined based on the dilutions.

2.3.3 Library screens. In-vivo excision of pBluescript phagemid from λZAP bacteriophage (Stratagene)

Two square plates (22x22cm) were preheated to 37°C and NZY top agar melted to 48°C. 150,000 pfu/µl of the library (1µl of the appropriate dilution) were added to 1.2 ml of XL1 Blue MRF' (OD_{600} of 0.5). The library reaction was incubated for 15 minutes at 37°C to allow the phage to attach to the cells. 60ml of NZY top agar were added and the mix plated onto the NZY agar plates (~30ml mix on each or 75,000 pfu per plate). The plates were incubated at 37°C for 6-8 hours. Once pinpoint size colonies have developed the plates were cooled down to 4°C to prevent the NZY top agar from sticking to the nylon membrane.

The subsequent steps of the library screen are described elsewhere. The transfer of the plaques to a nylon membrane (plaque lifts) is described in 2.5.13.3. The subsequent probe labelling and hybridisation are described in 2.5.13.1 and 2.5.13.2 respectively.

Typically three rounds of screens were carried out to isolate positive clones. The latter were plugged and transferred to eppendorf tubes containing 0.5ml phage buffer and 20µl chloroform. The construction of the λZAP-cDNA library allowed simple in-vivo excision and recircularisation of any cloned insert contained within the λ vector to form a pBluescript SK(-) phagemid containing the cloned cDNA insert. For the in-vivo excision a 250µl aliquot from the eluted phage was added to 200µl XL1-Blue MRF' cells (OD_{600}=1.0) with 1µl of ExAssist helper phage (concentration ~$1x10^6$pfu/µl) supplied by Stratagene. The mix was incubated for 15 minutes at 37°C to allow absorption of the phage. 3ml of LB were added and the mix incubated for 3 hours at 37°C with shaking (150rpm). The culture was then heated

for 20 minutes at 70°C and centrifuged at 1000 x g for 15 minutes. The supernatant, containing the excised phagemid, was collected and stored at 4°C. 10µl and 100µl of the phagemid stock were mixed with two 200µl aliquots of SLOR cells (OD_{600}=1.0). The tubes were incubated for 15 minutes at 37°C and 200µl of each plated on LB-ampicillin plates. The developing colonies contained double stranded pBluescript SK(-) phagemid with the cloned cDNA insert.

2.4 *Drosophila* methods

2.4.1 *Drosophila* measurements

The size of different *Drosophila* stages significantly varies. In many experiments, such as Developmental Northern and Western, the exact weight of tissue is needed for designing a meaningful and accurate experiment.

Table 2.4.1 *Drosophila* measurements

Developmental stage	Weight	Volume
1 Pair of ovaries	600µg	0.6µl
1 Testes	35ng (size: 0.1mm x 0.1mm x 0.5mm x 7)	35nl
1 Embryo	12µg (size: 0.175mm x 0.175mm x 0.5mm)	12nl
1 Third Instar larva	1.5-1.6mg	1.5µl
1 pupa	1.5mg	15µl
1 Male fly	700µg	0.7µl
1 Female fly	1.3mg	1.3µl

2.4.2 Maintenance of *Drosophila* stocks

All fly stocks were maintained on cornmeal food at 18°C. Some stocks with reduced viability were kept at 25°C (Table 2.4.2 (A)). In the event of mite infestations, filter paper strips soaked in 3% (v/v) Benzyl benzoate (in EtOH) were placed on the top of the fly food.

Genotypes of stocks used for mapping *didum* and in genetic and P-element screen experiments are given in the relevant chapters. Fly stocks for general use, like Wild type flies and balancers are given in Table 2.4.2 (B).

Table 2.4.2 (A) Fly food and related media

Food name	Components/Preparation	Final concentration
Fly food* (Staffen)**	25g Cornflour	2.5%
	50g Sugar	5%
	17.5g Lyophilised yeast	1.75%
	10g Agar	1%
	Dissolve in 1L of boiling water. Add:	
	0.5ml 10% Nipagin in 95% Ethanol	0.005%
Apple juice plates	5g Bacto-agar	2.5%
	5g Sucrose	2.5%
	150ml H_2O. Boil, then add:	
	50ml Apple juice	25%
	0.2ml 10% Nipagin in 95% Ethanol	0.01%
GFP plates	4g Agar	2%
	5g Sucrose	2.5%
	200ml H_2O. Boil, then add:	
	50ml Red grape juice	
	0.2ml 10% Nipagin in 95% Ethanol	0.01%

Table 2.4.1 (B) Common wild type and balancer fly stocks

Stock/ Synonym	Genotype	Chr	Comments/ Donor
Oregon R (+38, B)**	*Wild type*		Red eyed wild type, John Roote
w^{1118} (3605, B)	*Wild type*		White eyed wild type, Bloomington
5194 (CKG30, B)	*w*; L^2 Pin^1/CyO, P{w^{+mC}=GAL4-Kr.C}DC3, P{w^{+mC}=UAS-GFP.S65T}DC7*	1;2	2 Chr. GFP balancer, Thomas Kornberg
4533 (B)	*w*; In(2LR)$noc^{4L}$$Sco^{rv9R}$, b^1/CyO, P{w^{+mC}=ActGFP}JMR1*	1;2	2 Chr. GFP balancer, Michael Ashburner
TM3/TM6 (3720, B)	*y1 w*; TM3, Sb1/TM6B, Tb+*	1;3	3 Chr. balancer, David Cribbs

* "Chr" from Chromosome. Specifies the chromosome position of the insertion or mutation.

** In brackets are given the Bloomington Stock center numbers.

2.4.3 Collection of embryos

Embryos were collected on apple juice plates that were smeared with a small amount of yeast paste. Unless otherwise stated, the collection period was 18 hours. The embryos were washed off the plate with some water and collected onto a fine plastic sieve. The eggs were rinsed with 0.1% Triton X-100 to prevent them from sticking to the filter walls

2.4.3.1 Drosophila egg preparation for dark-field microscopy

Flies were left to lay eggs on apple juice plates for approximately six hours. The eggs were washed off the plate with some water and collected onto a fine plastic sieve. The eggs were collected into a clean eppendorf tube and rinsed with PBS. 15 to 20 eggs were placed into a drop of Hoyer's mounting media on a microscope slide and covered with a coverslip. The slides were then incubated at 65°C O/N and stored at RT.

2.4.3.2 Collection of GFP embryos and larvae. Fluorescent microscopy

GFP embryos (containing a GFP balancer chromosome) were collected on special non-fluorescent grape juice plates. The plates were analysed on a fluorescent, dissecting microscope (Leica) with the filter combination required for GFP observation (450-90/515). When necessary GFP and/or non-GFP embryos were selected and transferred into a drop of FISH mounting media and covered with a cover-slip. The slides were analysed at higher magnification and photographed on a fluorescent microscope fitted with a digital camera.

GFP embryos were left to hatch on the grape juice plates. First, second and third instar larvae were picked out of the plates and transferred into a drop of 4% p-Formaldehyde on a microscope slide and covered with a cover slip. The slides were incubated at RT for 15 min and the immobilised larvae photographed on a fluorescent microscope fitted with a digital camera.

2.4.4 Collection of late (third) instar larvae, pupae and sexed adults

Newly eclosed flies were placed in a freshly yeasted, 250ml fly bottle for 4 to 6 hours and then removed. The bottles were incubated at 26°C for the appropriate

time. After 96 hours the third instar larvae have crawled up the bottle walls. From there they were collected with a paintbrush.

Early pupae were picked out from the bottle walls after 120 hours of incubation. 5 to 15 hours later dark (late) pupae were collected.

Generally adult flies no older then 10 days were collected. The flies were anaesthetised with CO_2 (occasionally diethyl ether was used), sexed and placed in eppendorf tubes. Males and females were distinguished by identifying the specific sexually dimorphic characteristics such as: body size and shape, abdomen colour and the presence of sex combs in males.

2.4.5 Collection of virgin flies and crosses

Virgin flies were collected in one of the following ways. Generally the fly-bottles were emptied from any eclosed flies. Then every 6 hours (26°C) the newly eclosed flies were collected and sexed. Flies from the same sex were transferred together into vials with fresh flyfood.

When a limited number of pupae were available a different approach was undertaken. Dark pupae were picked out with a paintbrush and placed individually into vials with fresh flyfood. The newly eclosed flies were anaesthetised and sexed.

Generally, fly crosses were carried out by transferring 2 virgin females and 2-4 male flies into the same vial, containing fresh flyfood, supplemented with yeast paste. The flies were incubated at 18 or 26°C (depending on the genotype of the flies) for up to 4 days and then removed.

2.5 Nucleic acid methods

2.5.1 Common techniques for DNA and RNA purification

2.5.1.1 Phenol/Chloroform extraction

Protein impurities were removed from nucleic acid preparations by phenol/chloroform extraction. For DNA purification a phenol solution equilibrated to pH-8.0 is required; for RNA a phenol with pH-4.3 is required (see Table 2.1.3 (A)). An equal volume of phenol/chloroform was added to the DNA (RNA) solution, which was then mixed thoroughly by repeated gentle inversions for 5 minutes. The

suspension was centrifuged at 14,000 x g for 5 minutes and the top aqueous phase was transferred to a clean tube. This step was repeated until a clear interface between the phases was achieved. A final extraction with an equal volume of chloroform ensured complete removal of the phenol.

2.5.1.2 Precipitation of nucleic acids

DNA/RNA was precipitated from solution by adding 0.1 volume of 3M sodium acetate (pH-5.2) and 2.5 volume 99.7% EtOH. Alternatively DNA was precipitated with 0.6 volumes isopropanol. Typically the nucleic acid was precipitated for 1h at -20°C (no incubation was needed for the isopropanol precipitation) and centrifuged at 22,000 x g for 15 minutes at 4°C. The resulting pellet was washed once with 75% EtOH and centrifuged at 22,000 x g for 5 minutes at 4°C. The DNA pellet was air dried and resuspended in the appropriate volume of 10mM Tris-HCl, pH-8.0. The RNA was resuspended in RNase free water.

2.5.1.3 Estimation of the concentration of nucleic acids by UV spectrophotometry

UV spectrophotometry was used to determine the concentration and purity of nucleic acids. The DNA/RNA absorbance was measured at 260nm and 280nm using a two-beam Hitachi spectrophotometer. The concentration of the sample was calculated using the following formula:

$$\text{DNA/RNA Concentration in ng/}\mu\text{l} = \frac{A_{260} \text{ x DF x Sc}}{1000}$$

DF - Dilution Factor

Sc - Spectrophotometric conversion (Specific DNA/RNA absorption value)

Sc = 50 (double stranded DNA)

Sc = 40 (RNA or single stranded DNA)

The purity of DNA was estimated by assessing the ratio A_{260}/A_{260}: DNA was considered to be sufficiently free of protein at ratio greater than 1.8 (2.0 for RNA).

2.5.2 Genomic DNA preparation

Approximately 130mg of flies were ground in 250µl Fly Buffer using a plastic pestle. The volume was adjusted to 1ml with an additional 750µl of Fly buffer (Table 2.5.2). 8µl of RNase was added (80µg/ml final concentration) and the homogenate incubated for 30 minutes at 37°C. 60µl of Proteinase K was then added (620µg/ml final concentration), followed by a 30minute incubation at 50°C. Cell debris was pelleted by centrifugation (22,000 x g, 15 minutes, at 4°C). The supernatant was transferred to a clean eppendorf tube and extracted with phenol/chloroform.

The plasmid DNA was recovered from the solution by isopropanol precipitation. After a brief wash with 75% EtOH the DNA pellet was air dried at 42°C and resuspended in 70µl 10mM Tris-HCl, pH-8.0.

Table 2.5.2 Solutions required for genomic DNA preparation

Reactives/Chemicals	Composition/Preparation	Final concentration
Fly Buffer	5ml 1M Tris-HCl, pH-8.5	100mM
	800µl 5M NaCl	80mm
	2.5g Sucrose	5%
	2.5ml 10% SDS	0.5%
	5ml 0.5M EDTA, pH-8.0	50mM
	Adjust the volume to 50ml with ddH_2O	
Proteinase K	10mg Proteinase K	10mg/ml
	10µl 1M Tris-HCl, pH-7.5	10mM
	590µl ddH_2O	
	400µl Glycerol	40%

2.5.3 Plasmid DNA preparation

2.5.3.1 General purpose DNA miniprep

A 5 ml O/N culture was centrifuged to pellet the bacterial cells. The pellet was then resuspended by vortexing in 250µl Resuspension buffer. 350µl of Lysis solution was added and the mix was gently inverted several times. After a 5 minute incubation 350µl of Neutralising solution was added. The mix was inverted gently

four times and centrifuged at 22,000 x g for 15 minutes at 4°C. The supernatant was transferred to a clean eppendorf tube and the plasmid DNA recovered from the solution by isopropanol precipitation.

Table 2.5.3 Solutions required for plasmid DNA preparation

Reactives/Chemicals	Composition/Preparation	Final concentration
Resuspension buffer	500µl 1M Tris-HCl pH-7.6 200µl 0.5M EDTA pH-8.0 100µl 10mg/ml RNase (DNase free) Adjust the volume to 10ml with ddH_2O	50mM 10mM 100µg/ml
Lysis Solution	2ml 10N NaOH 10ml 10% SDS Adjust the volume to 100ml with ddH_2O	0.2M 1%
Neutralising solution	3.24g CH_3COOK (Mw: 98.14) Adjust the pH to 4.8 with Glacial Acetic acid (~2ml) Adjust the volume to 25ml with ddH_2O	1.32M

2.5.3.2 High purity plasmid miniprep

High quality DNA was required for sequencing and cloning experiments. QIAprep® Spin Miniprep kit (Qiagen) was used to isolate and purify for up to 20-30µg of plasmid DNA.

2.5.3.3 DNA midiprep

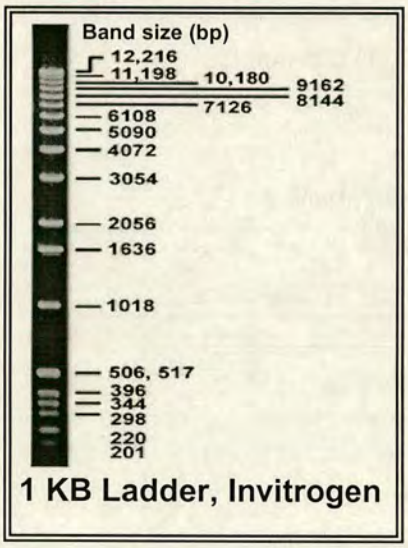
QIAfilter™ Plasmid midiprep kit from Qiagen was used to prepare up to 150µg of plasmid DNA following the supplier's instructions.

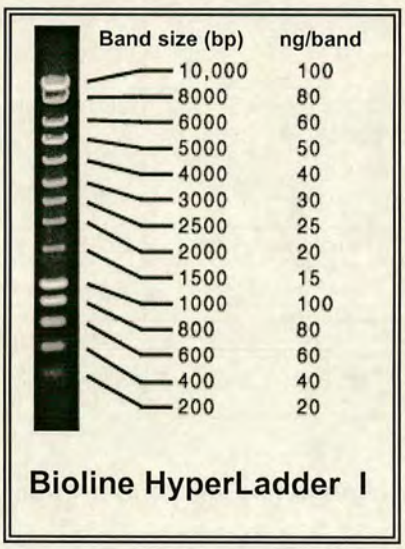
2.5.4 DNA agarose gel electrophoresis

Agarose gel electrophoresis was used to separate DNA molecules according to their size. 1.2% agarose gels were used for up to 8kb DNA molecules; for DNA molecules in the range 8-20kb 0.7 % gels were prepared. EtBr was added to the gel to a final concentration of 0.2mg/ml. The samples were premixed with loading buffer (Table 2.1.3) to a final concentration of 1x. The DNA samples were loaded on the

gel, and electrophoresed at 4-7 V/cm in 1xTAE buffer until the appropriate degree of separation was obtained.

Molecular standards were used to determine the size of the DNA fragments. Approximately 0.5µg of Invitrogen 1kb Ladder was applied per gel. To estimate the approximate amount of DNA in an unknown band Bioline HyperLadder was used. 5µl of the ladder would produce a set of bands with a known amount of DNA in each band. These were then compared to a particular fragment with an unknown amount of DNA.





Nucleic acids stained with EtBr were visualised on a long-wavelength UV transilluminator (365nm). A video imager was used to obtained permanent records of the gel pattern on thermal paper. These were subsequently scanned on desktop scanner and stored as JPG picture files.

2.5.5 Recovery of DNA fragments from agarose gel

QIAEX® II Gel Extraction Kit (Qiagen) was used to recover up to 15kb DNA fragments from agarose gels. The supplier's procedure was modified as follows: the DNA fragment was excised from the gel and dissolved in 3 volumes QX1 buffer with 2 volumes ddH_2O at 65°C for 5-10 minutes. 10µl of 3M Sodium acetate (pH-5.2) and 10µl QIAEX II resin were added. The sample was incubated for 10 minutes at RT with mixing every 2 minutes to keep the QIAEX II in suspension. The QIAEX II particles were pelleted at 14,000 x g for 1 minute and then washed once with QX1

buffer and two times with PE buffer (Qiagen). The pellet was air-dried at 37°C for 5-10 minutes. To elute the bound DNA 30µl of 10mM Tris-HCl, pH-8.0 were added to the pellet and the mix incubated for 5 minutes at 65°C. The sample was centrifuged at 22,000 x g and the cleared supernatant transferred into a new eppendorf tube. The elution step was repeated and the eluates combined.

2.5.6 DNA sequencing and sequence analysis

A typical (1x) sequencing reaction consisted of 8µl Labelling Mix (Rhodamine or Big DyeTM Terminator for Cycle Sequencing, ABI PRISM, Applied Biosystems), 3.2pmol of the appropriate primer, 50-200ng (PCR product) or 400ng (plasmid DNA) high quality DNA template and water to a final volume of 20 µl. The program for thermal cycling was as follows: 25 cycles, 95°C for 30 seconds, 50°C for 20 seconds and a 4 minute extension step at 60°C. To remove the unincorporated dye terminators the sequencing reaction was passed through a purification column. Either Edge Gel Filtration Cartridges (ver.11) from Edge BioSystems or genCLEAN columns supplied by Genpack were used for purification, following the manufacturer's instructions.

The amplified and labelled products were analysed on ABI PRISM 377 DNA sequencer (Perkin-Elmer Corporation) by Nicola Preston (ICMB-Edinburgh University) or Jill Lovell (ICAPB-Edinburgh University). Occasionally the products were analysed on ABI PRISM 3100 DNA Sequencer by Helen Williamson (ICMB-Edinburgh University).

2.5.7 Enzymatic reactions

2.5.7.1 Endonuclease restriction of DNA

Typically a 20µl reaction was carried out. When needed the standard reaction was scaled up proportionally. DNA sample (0.2-1.0µg), appropriate restriction buffer and 10U of the restriction enzyme were used per reaction. The reaction was incubated for 2 to 6 hours at the recommended temperature (usually 37°C) and then stopped by heat inactivation of the enzyme (20minutes at 65°C). The results of the restriction enzyme digest were analysed using agarose gel electrophoresis.

2.5.7.2 DNA phosphorylation

Phosphorylation of DNA is achieved by using T4 polynucleotide kinase and ATP. Common application of this reaction is a 5' DNA labelling and recovering the 5' Phosphate in dsDNA linkers to enable them to ligate with other DNA fragments. In general a 20μl reaction was set, containing 1μM ATP, 1x of the provided by the manufacturer buffer and 3-5U of the enzyme. The reaction was incubated for 30 minutes at 37°C.

2.5.7.3 DNA dephosphorylation

DNA dephosphorylation was carried out in 20μl reaction volume. The reaction contained the following components: the sample DNA, 1x of the supplied buffer, and 1-2U of alkaline phosphatase. For sticky end ligation, the mix was incubated for 30 minutes at 37°C and then another unit of the enzyme added. When ligating DNA fragments with blunt ends the incubation times were extended to 1 hour. Either CIAP or SAP alkaline phosphatases were used in the experiments. CIAP required the completed reaction to be stoped with 2μl of 0.5M EDTA and the DNA purified by phenol/chloroform and recovered by EtOH precipitation. SAP was simply inactivated by incubation at 65°C for 20minutes and the DNA precipitated with EtOH.

2.5.7.4 DNA ligation

DNA ligation was performed by incubating DNA fragments with appropriately linearised cloning vector in the presence of 1x buffer (with ATP) and 400U T4 DNA ligase (NEB) in 20μl reaction volume. The reaction was incubated O/N at 4°C when joining dsDNA with cohesive ends or 18°C for blunt ended DNA molecules. The T4 DNA ligase was then heat inactivated (20minutes at 65°C) and 1-2μl of the reaction used for transformation to competent cells.

2.5.8 DNA transformation to chemically competent cells

2.5.8.1 Preparation of competent cells

5ml O/N culture (usually XL1 Blue cells) was used to inoculate 200ml SOB^+ media (no antibiotic was added). The culture was grown at 37°C, 250rpm, until an

OD_{590} of 0.6 was reached (typically 5-6 hours). The cells were cooled down to 4°C and pelleted by centrifugation at 2500 x g for 10 minutes at 4°C. The pellet was resuspended gently in 14ml ice-cold TB buffer and 1ml DMSO added (7% final concentration). The cells were incubated for 10 minutes on ice, aliquoted in 100µls and snap frozen in liquid nitrogen. The competent cell were then stored at -80°C for up to 6 months.

2.5.8.2 Transformation

A 100µl aliquot of competent cells was thawed on ice for 5-10 minutes. Up to 50ng of plasmid DNA, in a volume not exceeding 4µl, was added to the aliquot. The reaction was incubated for 30 minutes on ice and then heat-shocked at 42°C for 60-90 seconds. After the cells were cooled briefly on ice, 600µl of SOC^+ media were added and the mix incubated at 37°C for 45 minutes. 200µl of the transformed cells were plated on appropriate selective agar plates.

2.5.9 Polymerase chain reaction (PCR)

The standard polymerase reaction consisted of the following components: 50pmol of the forward and reverse primer, 0.2mM of each dNTP (dATP, dCTP, dGTP, and TTP) supplied by Amersham Biosciences, 1x Taq Polymerase Buffer (Qiagen), 1x Q Solution (Qiagen) and 3U Taq Polymerase (Qiagen), approximately 50-100ng of DNA template and nuclease free water to a final volume of 50µl. The PCR reaction was carried out as follows: Initial denaturing step at 94°C for 2 minutes; then 30 cycles, step one at 94°C for 40 seconds, step two at 50-60°C (depending on the annealing temperature of the primers used) for 40 seconds, step three at 72°C for 30 seconds to 15 minutes (depending on the length of the amplified product, allowing 1 minute per 1kb). This was completed by a final 72°C elongation step for 10 minutes. Either a Touchdown Thermal Cycler (Hybaid) or Hybaid PCR Express Cycler (allowing a gradient PCR to be carried out) were used in the PCR experiments. Subsequently 5 to 10 µl of the total reaction were analysed by gel electrophoresis. Information about general sequencing primers is given in Table 2.5.8. The sequence of primers used in the analysis of *Br-C*, *didum* and the newly identified myosin genes in *Drosophila* are given in the relevant chapters.

Table 2.5.8 General purpose primers specific to pBluescript and other cloning vectors, used for sequencing and PCR experiments

Primer name	Sequence in 5' > 3' direction	Length nt	Tm °C
T3PR	AAT TAA CCC TCA CTA AAG GG	20	56
T7-20	TAA TAC GAC TCA CTA TAG GG	20	56
T7-LN	GTA ATA CGA CTC ACT ATA GGG	21	60
SP6	TAT TTA GGT GAC ACT ATA G	19	50
SP6-LN	ATT TAG GTG ACA CTA TAG AA	20	52
M13-17	GTA AAA CGA CGG CCA GT	17	52
M13REV-19	GGA AAC AGC TAT GAC CAT G	19	56
G70202	GGA ATT CTT TTT TTT TTT TTT TTT	24	37
Oligo dT	TTT TTT TTT TTT TTT	15	34
RP491	AAG CCC AAG GGT ATC GAC AAC	21	62
RP492	ATT GAA CTC GGC ACT GGC ACA	21	62

2.5.10 Reverse transcription and PCR (RT-PCR)

Reverse transcription was carried out in a 40µl reaction volume and contained the following: 4µg mRNA (total RNA), 1µM oligo (dT) primer, 10mM DTT, 1x RT buffer, 0.5mM of each dNTP (dATP, dCTP, dGTP, and TTP), and 200U Superscript II. Experiments were performed either with oligo $(dT)_{15}$ primer or PolyT-EcoRI primer (70202). The latter produced PCR products with higher quality due to its higher annealing temperature. The RNA and Oligo primer were premixed and heated to 70°C for 5 minutes and snap cooled on ice before adding the other reaction components. This step is needed to remove any secondary structures from the RNA template. The RT mix was incubated for 1 hour at 42°C and then stopped by heat inactivation of the enzyme at 70°C for 15 minutes. 2µl of the RT reaction was used as a template for subsequent PCR amplification. Standard PCR conditions were used as described in the previous section, 2.5.8.

2.5.11 Inverse PCR

Genomic DNA was prepared from flies containing P-element insertion(s) according to the protocol in section 2.5.2. Approximately 8µg of the gDNA was digested with Sau IIIA for 6 hours at 37°C. The reaction was stopped by heating the mix at 65°C for 20min. The digest was directly used in a 200µl ligation reaction with 800U of T4 DNA ligase (NEB). The reaction was incubated O/N at 4°C. The following day the DNA was recovered by EtOH precipitation and resuspended in 150µl 10mM Tris-HCl, pH-8.0. 5µl of the religated genomic DNA was used as a template for subsequent PCR amplification with P-element specific primers (Table 2.5.10). Standard PCR conditions were used as described in section, 2.5.8.

Table 2.5.10 Primers for Inverse PCR

Primer name	Sequence in 5' > 3' direction	Length nt	Tm °C
PLac1 (5'PZ)	CAC CCA AGG CTC TGC TCC CAC AAT	24	64
PLac4 (5'PZ)	ACT GTG CGT TAG GTC CTG TTC ATT GTT	27	65
SpLac2 (5'PZ)	GAA TTC ACT GGC CGT CGT TTT ACA A	25	64
Pry1(3'PZ)	CCT TAG CAT GTC CGT GGG GTT TGA AT	26	65
Pry4 (3'PZ)	CAA TCA TAT CGC TGT CTC ACT CA	24	62
W3-1 (3'lacW)	TGT CGG CGT CAT CAA CTC C	19	60
Spep1 (3'PZ)	GAC ACT CAG AAT ACT ATT C	19	52
Sp1 (5'PZ)	ACA CAA CCT TTC CTC TCA ACA A	22	62
Sp3 (3'PZ)	GAG TAC GCA AAG CTT TAA CTA TGT	24	62
Sp5 (3'lacW)	GCA TCA CAA AAA TCG ACG CTC AAG T	25	64
LACF (lacZ)	AGC TGG CGT AAT AGC GAA GA	20	60
LACS (lacZ)	CAC ACT TCG GCA CGT GAA TT	20	60

2.5.12 RNA preparation and electrophoresis

2.5.12.1 Methods for preparation of total RNA

TRIzol® Reagent (Life technologies) was used for general preparation of total RNA. The manufacturer protocol was changed slightly as follows: Approximately 50mg of tissue was homogenised in 300µl TRIzol for 1.5 minutes. A further 700µl TRIzol were added and the homogenate mixed by inverting. After a 5 minute

incubation at RT, 200µl of chloroform was added. The sample was vortexed for 15 seconds and then incubated at RT for 2 minutes. The phases were separated by centrifugation at 14,000 x g, for 15 minutes at 4°C. The top water phase, containing the RNA, was collected (~600µl) and mixed with 500µl of isopropanol. After 10 minutes at RT the RNA was pelleted by centrifugation at 14,000 x g, for 10 minutes at 4°C. The pellet was washed once with 75% EtOH and air-dried for 15-20 minutes at RT. The RNA was resuspended in 70µl RNase free water for 10 minutes at 65°C. Then the sample was centrifuged for 5 minutes at maximum speed (22,000 x g) and the cleared supernatant, containing the reconstituted RNA, transferred to a clean eppendorf tube. Usually up to 120µg of total RNA was isolated with this method.

Small amounts of total RNA (up to 50µg) were isolated with RNeasy Mini Kit (Qiagen), following the suppliers protocol. This was the preferred method when hard to obtain tissues like testes and ovaries were to be processed.

2.5.12.2 RNA gel electrophoresis

A 1.2% agarose gel was prepared in DEPC treated water. The gel was cooled down to 60°C. Then 20x MOPS (1x final concentration) and 37% formaldehyde (1.8% final concentration) was added and the gel allowed to solidify for 1 hour.

RNA samples (at least 10µg of total RNA or 1µg of mRNA per lane) were premixed with RNA sample buffer (Table 2.1.3) in a ratio 1:3. The probes were incubated at 65°C for 5 minutes, to remove any secondary structures from the RNA, and snap cooled on ice.

Gels were run in 1x MOPS buffer at 7 V/cm for 1-2 hours (check up gels) or O/N at 2-3 V/cm, for Northern blot experiments.

2.5.13 Hybridisation techniques

2.5.13.1 Radio labelling of DNA probes

Ready-To-Go™ DNA Labelling Beads (-dCTP) from Amersham Biosciences were used to label DNA probes with [^{32}P]dCTP (Amersham Biosciences), as instructed by the manufacturer. Approximately 100ng of DNA was used per labelling. Unincorporated nucleotides were removed by passing the probe through NICK™ Columns (Amersham Biosciences). These were preloaded with G-50

Sephadex for gel filtration allowing the high molecular weight DNA to pass faster through the column than the unincorporated nucleotides. The radio-labelled probes were denatured for 5 minutes in boiling water and snap cooled on ice.

2.5.13.2 Southern Blotting

The alkaline blotting technique was used for DNA transfer to Hybond™-N+ (Amersham Biosciences) nylon filter, following the manufacture's instructions. After gel electrophoresis DNA was depurinated in 0.25N HCl acid for 15 minutes (this step was not required when DNA fragments smaller than 10kb were processed). The transfer was carried out O/N in 0.4N NaOH. On the following day the filter was air dried and DNA fixed to the membrane by UV crosslinking (254nm + 365nm UV light) for 10 minutes.

2.5.13.3 Colony and plaque lifts

The cells or the bacterophage were plated in the usual way described in sections 2.2.3 and 2.3.1 respectively. The plates were precooled to prevent smearing of the colonies and separation of the top agar in the phage plates. The transfer was performed by placing a nylon membrane on the surface of the agar for 2 minutes. The membrane was stabbed asymmetrically with a needle to ensure correct orientation of the colonies/plaques in the subsequent manipulations. The membranes were denatured for 2 minutes in Denaturing buffer, then placed for 2 minutes in Neutralising buffer and washed for 30sec in Rinse buffer. The nylon filters were air-dried and the DNA fixed by UV crosslinking (254nm + 365nm UV light) for 10 minutes.

Table 2.5.13.3 Solutions required for colony and plaque lifts

Solution	Components/Preparation	Final concentration
Denaturing buffer	150ml 5M NaCl 25ml 10N NaOH 325ml H_2O	1.5M 0.5M
Neutralising buffer	150ml 5M NaCl 250ml 1M Tris-HCl, pH-8.0 100ml H_2O	1.5M 0.5M
Rinse buffer	50ml 20x SSC 100ml 1M Tris-HCl, pH-7.5 350ml H_2O	2x 0.2M

2.5.13.4 Northern Blot

The transfer of RNA from agarose gel to nylon filter (Hybond™-N+) was similar to the DNA transfer, with the following differences: After completion of the electrophoresis the RNA gel was pre-soaked in 10X SSC for 30 minutes at RT. The transfer was allowed to proceed O/N in 10x SSC. On the following day the filter was air-dried and the RNA fixed to the membrane by UV crosslinking (254nm + 365nm UV light) for 5 minutes.

2.5.13.5 Hybridisation with [^{32}P]dCTP radio-labelled probes

Filters were prehybridised in pre-hybridisation solution (DNA hybridisation buffer without the probe) at 65°C for 6 hours to O/N. The denatured radio-labelled probe was added to the prehybridisation buffer and the membranes incubated at 65°C O/N. The day after the membranes were washed as follows: 15 minute low stringency wash at 65°C in 2x SSPE, 0.1% SDS; 20 minute medium stringency wash at 65°C in 1x SSPE, 0.1% SDS, and 15 minute high stringency wash at 65°C in 0.1x SSPE, 0.1% SDS.

RNA blots were prehybridised and hybridised in a similar way, to DNA blots, but in RNA hybridisation buffer at 45°C. The membranes were washed two times for 20 minutes at 45°C in 1x SSC, 0.1% SDS (medium stringency wash) and once in 0.1x SSC, 0.1% SDS at 45°C (high stringency wash). The filters were then sealed in polythene bags, placed in a autoradiography cassette with intensifying screens, and

exposed to blue sensitive film (Kodak) at -70°C. The films were developed in an X-ograph CompactX2 automatic film processor.

2.5.13.6 Removal of probe and re-use of blots

Southern and Northern blots were stripped by placing them in boiling 0.1% SDS and allowing it to cool to RT. Films were autoradiographed to check that the probe had been removed.

2.5.14 *In-situ* hybridisation to mRNA

2.5.14.1 Preparation of digoxigenin-labelled probes

- Single-stranded digoxigenin-labelled DNA probes were generated by using cloned cDNA as a template. To avoid any enzymatic restrictions a PCR was performed with primers surrounding the MCS, thus amplifying only the cDNA, excluding the vector sequence. Then a second round PCR was carried out using DIG DNA labelling mix (Roche Molecular Biochemicals) and specific primer to generate ssDNA (antisense for the probe or sense for the negative control). 5µl of the probe was checked on agarose gel. Prior to hybridisation the digoxigenin-labelled probe was boiled for 30 minutes to remove any secondary structures and break down the DNA to shorter, easy to penetrate fragments.
- Single-stranded digoxigenin-labelled RNA probes were generated by *in vitro* transcription, using the Roche Molecular Biochemicals DIG RNA labelling kit, as instructed by the accompanying manual.

2.5.14.2 In-situ hybridisation to whole-mount ovaries using DNA probes

The protocol is based on the procedure previously described (Tautz and Pfeifle, 1998) and modified as follows. The ovaries were collected in Ringer's saline and fixed for 20 minutes in PMS at RT. After rinsing in PBT (3 times, 5 minutes each) the ovaries were digested in 10µg/ml Proteinase K for 20 minutes at RT. The digest was stopped by washing in PBT containing 2mg/ml Glycine. This was followed by three 5 minute washes in PBT. The ovaries were re-fixed for 10 minutes in PMS and then washed 3 times (5 minutes each time) in PBT. Prehybridisation was

carried out for 1 hour at 45°C in 500µl DNA Hybrix. The ovaries were hybridised overnight at 45°C in DNA Hybrix containing digoxigenin labelled probe (1:20 dilution).

After the hybridisation the ovaries were washed with preheated to 45°C PBT/DNA Hybrix for 20 minutes. This was followed by three washes, 10 minutes each in PBT. For detection a 1:1000 dilution of anti DIG-AP conjugated antibody was used (Table 2.6.2.3 (A)). The anti DIG-AP antibody was preabsorbed with postfixed wild type (Oregon R) ovaries at 4°C overnight.

The staining reaction was performed in Staining buffer containing 5mM Levamisole, 4.5µl/ml NBT and 3.5µl/ml X-Phosphate (Table 2.1.3 (C)). The colour reaction was stopped by washing the ovaries in PBS. The ovaries were mounted in a mixture of Glycerol/PBS media and analysed on a Nomarski microscope.

2.5.14.3 In-situ hybridisation to whole-mount ovaries using RNA probes

The *in-situ* hybridisation using DIG-labelled RNA probes is similar to the hybridisation with DIG-labelled DNA probes. RNA Hybrix was used instead of DNA Hybrix and the hybridisation was carried out at 65°C. This method produced very strong staining with insignificant background. Since it proved to be very sequence specific, it was only used with probes that have shown better staining with this procedure than with the hybridisation utilising DNA labelled probes.

2.5.15 Nuclear Hoechst staining

Ovaries were dissected in *Drosophila* Ringer's and fixed in FBS buffer for 20 minutes. This was followed by two 10 minute washes in PBT. The ovaries were stained for 5 minutes in 1µg/ml Hoechst 33258 dissolved in PBS. Residual Hoechst was removed by two PBT washes, 20 minutes each. The ovaries were mounted in PBS/Glycerol (1:4) and examined under a fluorescent microscope.

2.5.16 BrdU labelling

Ovaries were dissected at room temperature in 1x Grace's medium (Flow laboratories) and incubated for 1 hour in 15µM BrdU (Sigma) in Grace's medium (Lilly and Spradling, 1996). After washing in EBR the ovaries were fixed for 20

minutes in FBBW Buffer (Lin and Spradling, 1993), followed by 1 hour denaturing in 2N HCl and 2 minute neutralisation in 100mM Sodium tetraborate. The tissue was rinsed several times in PBT and blocked for 1 hour in 5% NGS in PBX (0.5% Triton X-100 in 1x PBS). After an O/N incubation in 1:20 dilution of the anti-BrdU antibody (Becton-Dickinson) a detection was carried out with HRP-conjugated secondary antibody (1:25 dilution). Sigma Fast™ DAB peroxidase substrate was used in the peroxidase colour reaction. The latter was enhanced with 20µl 1M Ni SO_4 per 1 ml staining solution.

2.6 *Protein techniques*

2.6.1 Protein sample preparation

2.6.1.1 Protein preparation from Drosophila tissues

Protein samples from different developmental stages of *Drosophila* were prepared in the following way (Table 2.6.1). Tissues were dissected in *Drosophila* Ringer's and homogenised in the appropriate volume of Protein sample buffer (Table 2.1.3 (B)). Samples were boiled for 5 minutes and cooled on ice. The cell debris and body particles were removed by centrifugation at 22,000 x g for 5 minutes and the cleared supernatant transferred to a clean eppendorf tube. The samples were stored at -80°C.

Table 2.6.1 Protein probe preparation

Developmental stage	Protein sample buffer
3 Ovaries	110µl
30 Testes	100µl
40 Embryos	60µl
2 Third Instar larvae	150µl
2 Light pupae (LP)	150µl
2 Dark pupae (DP)	150µl
3 ♂ flies (carcass)	110µl
2 ♀ flies (carcass)	150µl

2.6.1.2 Protein preparation of cell extracts

Proteins samples from bacterial cultures were prepared by centrifuging 1ml of O/N culture (OD_{600}= 2.0) at 14,000 x g for 1 minute. The pellet was resuspended in 100µl of Protein sample buffer, vortexed and the suspension boiled for 5 minutes. The cell debris was removed by centrifugation at 22,000 x g for 5 minutes and the cleared supernatant transferred to a clean eppendorf tube. The samples were stored at -80°C.

A solution of the recombinant protein was prepared by pelleting down 3ml of induced BL21DE3 Lys cells (OD_{600}-1.6) containing the recombinant expression vector, pBMF33 (the expression vector was produced by Bryce McIver, personal communications). The pellet was resuspended in 1ml PBS, mixed with 20µl Lysozyme from a Stock and incubated on ice for 40 min. The suspension was then sonicated for 2 minutes at maximum power and centrifuged to remove the cell debris. As a control in the western blots we used protein extract prepared from the sonication of induced BL21DE3 cells that contain only the cloning vector pRSET-C before the subcloning of myosin V tail sequence.

2.6.1.3 Protein preparation from inclusion bodies

Cells containing the recombinant construct were induced with 1mM IPTG for 4 hours to express the recombinant protein. The bacterial culture (usually 500ml) was centrifuged at 10,000 x g for 15 minutes. The cell pellet was resuspended in 3ml X buffer (Table 2.6.6) and stored at -20°C for up to two weeks.

Next, the X buffer was removed by centrifugation at 10,000 x g for 10 minutes and the cells resuspended in X1 Buffer (3ml X1 buffer per 1g cells). 80µl of Lysozyme stock was added to the cells and the solution incubated for 40 minutes at 4°C. The cells were placed into an ice bucket and sonicated for 3 minutes at full power. The suspension was centrifuged at 10,000 x g for 15 minutes and the pellet washed in 5ml X1 Buffer. After a second centrifugation at 10,000 x g for 15 minutes the pellet was resuspended in X1U buffer (4.5ml per 1g cells) and incubated O/N at 4°C. On the following day the solution was centrifuged for 40 minutes at 10,000 x g. The supernatant containing the solubilised protein was collected and stored at -80°C.

2.6.2 Bradford assay

The Bradford assay was used to determine the concentration of the recombinant protein. The methods is based on the proportional binding of the dye Coomassie to proteins. The assay is colorimetric; as the protein concentration increases, the colour of the tested sample becomes darker. The protein concentration of a test sample is determined by comparison to that of a series of protein standards known to show a linear absorbance profile in this assay. To build a standard curve two stock solutions of BSA (Bovine Serum Albumin) were prepared, 25mg/ml and 50mg/ml. They were diluted accordingly to obtained a set of working solutions with known protein concentration (Table 2.6.2A).

Table 2.6.2A Preparation of test samples for Bradford protein assay

Protein standard mg/ml	Sample volume µl	Added 1xPBS µl	Absorb-ance I	Absorb-ance II	Average absorb-ance
0.1	4 (25mg/ml stock)	996	0.115	0.123	0.119
0.2	8 (25mg/ml stock)	992	0.231	0.232	0.232
0.4	16 (25mg/ml stock)	984	0.404	0.433	0.418
0.6	24 (25mg/ml stock)	976	0.574	0.603	0.588
0.8	32 (25mg/ml stock)	968	0.710	0.726	0.718
1.0	40 (25mg/ml stock)	960	0.771	0.809	0.790
1.2	24 (50mg/ml stock)	976	0.890	0.963	0.926
1.4	28 (50mg/ml stock)	972	0.950	0.973	0.960

Each sample was prepared by mixing 50µl of protein standard with 2.5ml of Bradford reagent. The samples were incubated for 10 minutes at room temperature. The absorbance was measured spectrophotometrically at 595nm and a standard curve was built based on the obtained absorbance values.

To determine the concentration of the recombinant protein 50µl of the protein solution was diluted in 350µl 1xPBS buffer (eight fold dilution) and 50µl of this solution was diluted again in 350µl PBS (64x dilution). Three samples of each dilution were made by mixing 50µl of the protein with 2.5ml Bradford reagent. Two measurements were taken from each sample (in Table 2.6.2B are shown the averaged values).

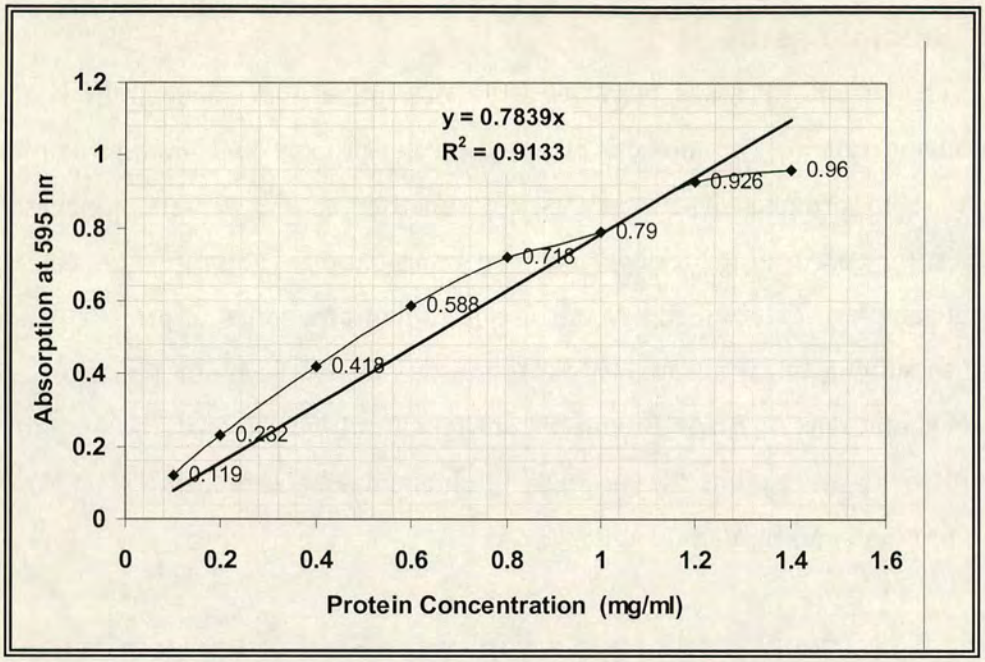


Table 2.6.2B Results from the spectrophotometrical measurments

Sample dilution	Absorbance of sample I	Absorbance of sample II	Absorbance of sample III	Average absorbance	Protein concentration
8x	0.241	0.259	0.261	0.253	2.0 mg/ml
64x	0.036	0.039	0.037	0.037	2.4 mg/ml

The mean value obtained from the measurements of the two protein samples is 2.2mg/ml. Thus the amount of recombinant protein recovered from the two affinity purifications was estimated to be 4.4 mg based on spectrophotometrical analysis. Approximately half of this protein was used to generate antibody in sheep.

Solution	Components/Preparation
Bradford Reagent	100mg Coomassie Brilliant Blue G-250 50ml Et-OH (95%) 100ml Phosphoric acid (85% w/v) Adjust the volume to 1 litre

2.6.3 SDS-Polyacrilamide gel electrophoresis (PAGE)

The PAGE glass plates were set according to the manufacturer's instructions. Approximately 20ml of Separation mix (Table 2.6.1) was poured between the plates and overlaid with 2ml H_2O to even the surface. The gel was allowed to set for 20 minutes at RT. Then the water was decanted and 5ml of Stacking mix was poured on the top of the separation gel. After polymerisation the gel assembly was fixed to the electrophoresis tank and 1x TGS buffer added. Protein samples were loaded and the gel run at 40V O/N or at 150 V for 4 hours.

Table 2.6.1 Solutions required for SDS-PAGE

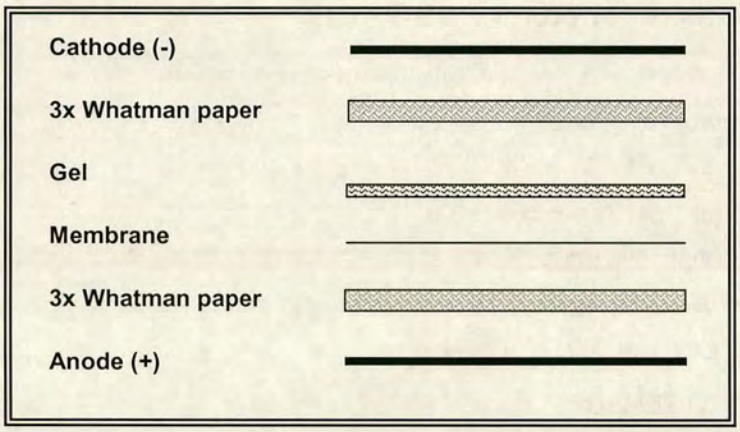
Solution	Components/Preparation
Separating mix	20ml 10% (for 21-100kDa proteins) 7.9ml ddH_2O 5ml 1.5M Tris-HCl, pH-8.8 200µl 10% SDS 6.7ml 30% Protogel [30%Acrylamide/0.8%Bisacrylamide (37.5:1)] 200µl 10% (w/v) $(NH_4)_2S_2O_8$ 20µl TEMED
Stacking mix	3.44ml ddH_2O 630µl 1M Tris-HCl, pH 6.8 50µl 10% SDS 830µl 30% Protogel [30%Acrylamide/0.8%Bisacrylamide (37.5:1)] 50µl 10% (w/v) $(NH_4)_2S_2O_8$ 10µl TEMED

2.6.4 Western Blot and immunodetection

2.6.4.1 Semi-dry transfer

Proteins were transferred from a polyacrylamide gel to Hybond™-C super nitrocellulose membrane (Amersham Biosciences) by semi-dry transfer. The transfer assembly was set according to the manufacturer's instructions. The stacker from the ready protein gel was removed and the gel placed on the top of pre-cut to the same size membrane. The transfer was run at 400mA (approximately 15V) for 2 hours.

Assembly sandwich



2.6.4.2 Ponceau S staining

To check the efficiency of the transfer the nitrocellulose membrane was stained with Ponceau S for 10 minutes. The excess stain was removed by briefly rinsing in H_2O.

2.6.4.3 ECL immunodetection

The membrane was incubated in 50ml Blocking buffer O/N at 4°C. Alternatively the incubation was carried out for 1 hour at RT with shaking (40rpm). The membrane was transferred to a nylon bag containing 16ml 1x TBS and the Primary antibody added at the appropriate dilution (Table 2.6.2.3 (A)).

Table 2.6.2.3 (A) List of primary antibodies

Antigen	Host animal	Working concentration		Source, Product number
		Immuno Histochemistry	Western Detection	
Br-C (Z1, Z3, and Core)	Mouse	1:50-100	n/a	Hodgetts (1995)
BrdU (mono*)	Mouse	1:20	n/a	Becton Dickinson No: 347580
Calmodulin (mono*)-Cy3	Mouse	1:25-50	n/a	Sigma C-7055
Digoxigenin -AP	Sheep	1:1000	1:10000	Roche No: 1 093 274
Myosin VA *(Drosophila)*	Sheep	1:50	1:4000	Bownes lab
Myosin VS *(Drosophila)*	Sheep	1:10	1:2000	Bownes lab
Non-muscle myosin II (*Drosophila*)	Rabbit	1:100	1:4000	Kiehart and Fegali (1986) No: 656 (RAM #V, pg015, 4/16/86
Non-muscle myosin II (vertebrate)	Rabbit	1:20	n/a	Biogenesis: No: 6490-1004

* "mono" from monoclonal antibodies

** "S" stays for non-purified serum; and "A" for affinity purified antibodies

After O/N incubation at 4°C the membrane was washed three times, 10 minutes each, in TBST with shaking at 40rpm. Then the membrane was blocked for 20 minutes at RT in 0.4% Normal serum from the animal species in which the secondary antibody has been generated. Subsequently the membrane was placed in a nylon bag containing 16ml 1x TBS with the appropriate concentration of Secondary Antibody (Table 2.6.2.3 B).

Table 2.6.2.3 (B) List of Secondary antibodies

Antigen	Host	Labelling	Working concentration		Source, Product number
			Immuno Histochemistry	Western Detection	
Mouse IgG	Goat	HRP	1:100	1:4000	Promega W4021
Mouse IgG	Goat	TRITC	1:80	n/a	Sigma T7782
Rabbit IgG	Goat	HRP	n/a	1:4000	Sigma A8275
Rabbit IgG	Goat	FITC	1:100	n/a	Sigma F9887
Sheep IgG	Donkey	HRP	n/a	1:4000	Sigma A3415
Sheep IgG	Donkey	FITC	1:100	n/a	Sigma F7634
Sheep IgG	Donkey	Cy2	1:50	n/a	Jackson Immuno Research No: 713-225-147

After 1 hour incubation at RT the membrane was washed three times, 10 minutes each, in TBST with shaking at 40rpm. Then the membrane was immersed into 20ml ECL Reagent with 8µl 30% H_2O_2 and incubated for 1 minute (Table 2.6.2.4). The membrane was then sealed in a polythene bag, placed in an autoradiography cassette with intensifying screens, and exposed to blue sensitive film (Kodak) for 1 to 10 minutes at RT.

2.6.4.4 Membrane Stripping

The membrane was submerged in stripping buffer (Table 2.6.2.4) and incubated at 50°C or 70°C (for higher stringent conditions) for 30 minutes with agitation.

Table 2.6.2.4 Solutions required for ECL detection

Solution	Components/Preparation	Final concentration
Blocking Buffer	15g Casein 15ml Ethanol Add 50ml 1x TBS. Mix well. Add additional 100ml 1x TBS and make sure that the casein is completely dissolved. Add 135 ml 1x TBS	5%
Stripping buffer	3.47ml 14.4M β-Mercaptoethanol (Mw: 78.01) 100ml 10% SDS 31.25ml 1M Tris-HCl, pH-6.7	100mM 2% 62.5mM

2.6.5 Immunohystochemical detection in whole mount ovaries

Ovaries were dissected from yeasted for two days flies in Ringer's solution. The anterior parts of the ovaries were torn apart to facilitate antibody penetration. The ovaries were transferred to an eppendorf tube containing FBS and fixed for 20 minutes at room temperature. The fixative was carefully removed and the ovaries washed twice in PBS. Endogenous peroxidases were destroyed by treating with 1% H_2O_2 for 5-10 min (this step was required only when HRP-conjugated secondary antibody was used). The ovaries were washed again in PBS for 10 minutes and blocked in 5% NGS in PBX (0.5% Triton X-100 in 1x PBS) for 1 hour at RT (The blocking agent should always be the normal serum from the animal species in which the secondary antibody was generated). Primary antibody was added at the appropriate dilution and the ovaries incubated O/N at 4°C. Residual antibody was washed away with three changes of PBT with 1 hour incubation per change. The ovaries were blocked for 30 min in 5% NGS in PBX. Secondary Ab (antibody) was then added at the appropriate dilution and the ovaries incubated for 1 hour at RT. Excess secondary antibody was removed with three PBT washes at 30 minutes intervals.

When a HRP-conjugated secondary antibody was used for detection, DAB staining solution was added and the ovaries allowed to stain for 10-30 minutes before washing with PBS to stop the reaction. The DAB colour reaction was enhanced with

the addition of 20µl 1M $NiSO_4$ per 1 ml staining solution. Stained ovaries were mounted in PBS/Glycerol (1:4) and analysed on a Nomarski microscope.

Fluorescently stained ovaries (fluorescent secondary antibody was used for detection) were analysed and photographed on a fluorescent microscope fitted with a digital camera.

2.6.6 β-Galactosidase staining

Ovaries were dissected in *Drosophila* Ringer's and stained at room temperature for 3 hours to O/N in Galactosidase Staining buffer (Table 2.1.3) supplemented with 0.2% X-Gal. On the following day the ovaries were rinsed in PBS and mounted in PBS/Glycerol (1:4) for microscope analysis.

2.6.7 F-actin staining

Ovaries were dissected in *Drosophila* Ringer's and fixed in FBS for 20 minutes. The tissue was permeabilised for 10min in PBX (0.5% Triton X-100 in 1xPBS), and then stained for 30min in 6µM solution of Rhodamine-conjugated phalloidin at RT. The stained ovaries were washed twice for 10 min in PBS and mounted in FISH media for fluorescent microscopy.

2.6.8 Purification of Histidine-tagged proteins on Ni-Agarose column

2.6.8.1 Column preparation

To purify the protein we used Ni-NTA-Agarose. This resin allows the purification of His-tagged fusion proteins. In general the recombinant protein is adsorbed on a column prefilled with Ni-NTA-Agarose. The column is then washed and the protein eluted under high salt conditions.

2.5ml (drained volume) of Ni-NTA-Agarose (Qiagen) was placed in 10ml syringe, plugged with glass wool. The column was washed with 20ml X1 buffer and left in storage overlaid with 10ml X1 buffer.

2.6.8.2 Affinity purification

The Ni-NTA-Agarose column was washed once with 10ml X1 buffer and then equilibrated with 10ml X1U buffer. The protein solution in X1U buffer was passed three times through the column. 15ml of Wash buffer was passed through the column to remove the unbound proteins. The recombinant protein was eluted with 10ml Elution buffer. Ten 1ml fractions were collected and analysed on a PAGE (polyacrylamide) gel. The protein was usually in fractions from 2 to 5.

2.6.8.3 Protein dialysis

The fractions containing the recombinant protein were combined (approximately 15ml in total) and dialysed against 2L 1x PBS, O/N. The protein was concentrated by ultracentrifugation using UFV4BTK25, Ultrafree®-4 Centrifugal Filter unit from Millipore.

Table 2.6.6 List of solutions required for protein purification

Buffer	Components/Preparation	Final concentration
Imidazol, pH-8.0	1.36g Imidazol (Mw: 68.08) Adjust pH to 8.0 with cHCl Adjust the volume to 20ml	1M
Lysozyme	50mg Lysozyme 1ml ddH_2O	50mg/ml
X Buffer	5ml 1M Tris-HCl, pH-8.0 0.5ml 1M Imidazol, pH-8.0 50ml 5M NaCl 0.5ml Tween 20 50ml Glycerol Adjust the volume to 500ml with ddH_2O	10mM 1mM 500mM 0.1% 10%
X1 Buffer	5ml 1M Tris-HCl, pH-8.0 1ml 1M Imidazol, pH-8.0 50ml 5M NaCl 0.5ml Tween 20 Adjust the volume to 500ml with ddH_2O	10mM 2mM 500mM 0.1%
X1U Buffer	5ml 1M Tris-HCl, pH-8.0 1ml 1M Imidazol, pH-8.0 50ml 5M NaCl 0.5ml Tween 20 36g Urea (Mw: 60.06) Adjust the volume to 500ml with ddH_2O	10mM 2mM 500mM 0.1% 6M
Wash Buffer	1ml 1M Tris-HCl, pH-8.0 1ml 1M Imidazol, pH-8.0 10ml 5M NaCl 1ml 10% Tween 20 36g Urea (Mw: 60.06) Adjust the volume to 100ml with ddH_2O	10mM 10mM 500mM 0.1% 6M
Elution Buffer	1ml 1M Tris-HCl, pH-8.0 1.36g Imidazol (adjust the pH to 8.0 with c. HCl) 10ml 5M NaCl 1ml 10% Tween 20 36g Urea (Mw: 60.06) Adjust the volume to 100ml with ddH_2O	10mM 200mM 500mM 0.1% 6M

2.6.9 Purification of antibodies

2.6.9.1 Caprylic acid precipitation

Caprylic acid, at the right concentration and at a particular pH, will precipitate most serum proteins, except IgG. To purify the IgG fraction, 50ml of polyclonal antiserum was transferred to a suitable container with 2 volumes of 60mM Sodium Acetate pH-4.0 (Table 2.6.7). The solution was set on a magnetic stirrer and 3.75ml caprylic acid slowly added. This was achieved by setting up a disposable syringe barrel, fitted with a 19-G needle, over the beaker and adding the caprylic acid to the syringe barrel. This was followed by a 30 minute incubation on the magnetic stirrer. The precipitate was removed by centrifugation at 2000 x g for 20 minutes at RT. The supernatant was filtered through grade 4 paper to remove the fines and dialysed against 1x PBS. Subsequently the immunoglobulins were concentrated by ultracentrifugation using UFV4BTK25, Ultrafree®-4 Centrifugal Filter unit from Millipore.

2.6.9.2 Affinity purification of antibodies

Antibodies against HIS-tagged proteins were purified by binding to recombinant protein already attached to Ni-NTA-Agarose (Qiagen) and then eluting them with low or high pH buffer.

Ni-NTA-Agarose column was prepared and recombinant protein attached to the agarose as described in section 2.6.6. The column was then washed with 20ml PBS. 5ml of the caprylic acid purified antibody was diluted with 10ml PBS and passed three times through the column. This was followed by a column wash: once with 10ml PBS and once with 10ml High salt buffer. The antibodies were eluted with 10ml Glycine, pH-2.5. Ten 1ml fractions were collected in eppendorf tubes each containing 100µl 1M Tris-HCl, pH-8.0. The column was washed with 10ml 50mM Tris-HCl, pH-8.8. A high pH elution was carried out with 10ml Triethylamine, pH-11.5. Ten 1ml fractions were collected in eppendorf tubes each containing 100µl 1M Tris-HCl, pH-7.4. Next, the column was washed twice: once with 10ml 50mM Tris-HCl, pH-8.8, and once with 10ml PBS. The column was stored in 20% EtOH in PBS at 4°C.

The collected fractions were analysed by PAGE. The fractions containing the purified antibodies were combined and dialysed against 2 litres of PBS, O/N. The antibodies were concentrated by ultracentrifugation using UFV4BTK25, Ultrafree®-4 Centrifugal Filter (Millipore).

Table 2.6.7 List of solutions required for antibody purification

Buffer	Components/Preparation	Final concentration
Na Acetate, pH-4.0	4.08g $CH_3COONa.\ 3H_2O$ (Mw:136.08) Adjust the pH to 4 with Glacial Acetic acid Adjust the volume to 500ml with ddH_2O	60mM
High salt buffer	2ml 1M Tris-HCl, pH-7.5 10ml 5M NaCl Adjust the volume to 100ml with ddH_2O	20mM 500mM
Glycine, pH-2.5	3.75g Glycine (Mw: 75.07) Adjust pH to 2.5 with cHCl Adjust the volume to 500ml with ddH_2O	100mM
Triethylamine, pH-11.5	0.506g Triethylamine (Mw: 101.19) Adjust pH to 11.5 with cHCl Adjust the volume to 100ml with ddH_2O	50mM

Chapter Three: Molecular characterisation of *Drosophila* Myosin V

3.1 *Introduction*

3.1.1 Myosin motors

Myosins constitute a superfamily of motor proteins that convert energy from ATP hydrolysis into mechanical movement along actin filaments (Hasson and Mooseker 1995; Mooseker and Cheney 1995). Phylogenetic analysis currently places myosins into 18 classes based on class specific features of their conserved motor domain (for detailed discussion on the myosin superfamily see Chapter Six) (Sellers 2000; Berg et al. 2001). Traditionally the myosins have been divided into two classes depending on whether they form monomers or dimers (Fig. 3.1.1). The conventional myosin of muscle and non-muscle cells form class II myosins. They are complex molecules of four light chains bound to two heavy chains, that form bipolar filaments via interactions between their coiled-coil tails (type II). Class I myosins are smaller monomeric myosins referred to as unconventional myosins.

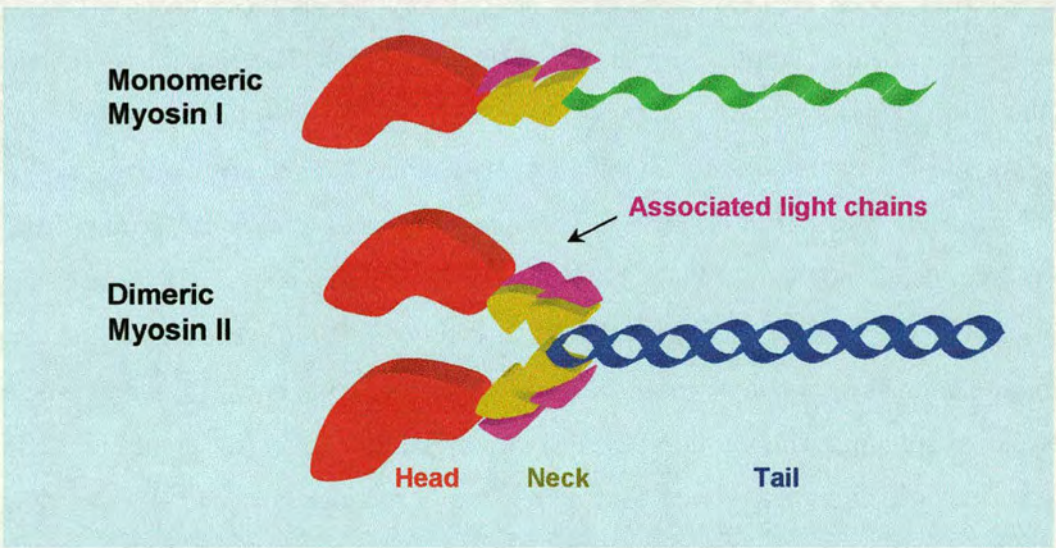


Figure 3.1.1 Model for the structure of myosin molecules

Each myosin molecule is composed of three main domains: a motor head domain, a regulatory neck domain and a tail domain. The head domain is highly conserved among all myosins. The neck domain may consist of a variable number of IQ motifs that bind light chains. The tail is the most divergent part in each myosin. The presence of coiled-coil motifs in the tail determines whether the molecule will be a dimer or not.

Now there are at least 16 other classes of unconventional myosins (Berg et al. 2001). Recent studies have shown that myosins are involved in a variety of cellular functions such as membrane trafficking, cell movements, maintenance of the cell architecture and signal transduction (Wu et al. 2000; Mermall et al. 1998; Berg et al. 2001).

All known myosins comprise an N-terminal head domain, a neck regulatory domain and a class specific tail domain (Fig.3.1.1) (Mooseker and Cheney 1995; Berg et al. 2001). The motor head domain (approximately 80kDa) is highly conserved among the myosins of the same class and amongst various classes of myosins. It contains ATP and actin binding sites and is responsible for the mechanochemical properties of the protein (Gilbert and Mackey 2000). Myosins show an actin stimulated Mg^{2+}-ATPase activity, thus converting the energy stored in ATP into mechanical force (Volkmann and Hanein 2000). The latter is used to move the myosin molecules along actin filaments or to translocate other molecules (Hasson and Mooseker 1995; Langford 1995).

The neck domain contains regulatory sites composed of IQ (isoleucine/glutamine) motifs, repeats of 23-30 amino acids. The consensus sequence of the IQ motifs is IQxxxRGxxxRK (Mercer et al. 1991). Each IQ motif provides a binding site for a calmodulin, a small 17 kDa subunit or related proteins of the EF-hand family (Kawasaki et al. 1998; Khan and Komiyama 2001). EF proteins have helix-loop-helix motifs in which the loop contains highly conserved residues that bind Ca^{2+} ions (Maki et al. 1997; Ikura 1996; Niki et al. 1996). The size of the neck domain varies from one to seven IQ repeats. Additionally the neck is often the site of alternative splicing. This produces necks with variable lengths (different number of IQ motifs), which are associated with its regulatory function. In general, calmodulin activates a diverse group of target cellular proteins when bound to Ca^{2+} including the myosins. Since each calmodulin molecule binds four Ca^{2+} ions and binding each Ca^{2+} facilitates binding of additional Ca^{2+} a small change in the concentration of Ca^{2+} leads to a large change in the level of active calmodulin. This in turn leads to a rapid change in the myosin activity. Interestingly most of the unconventional myosins contain IQ motifs that bind calmodulin with the highest affinity in the absence of Ca^{2+}. The regulation may proceed either as a direct consequence of

association/dissociation of the calmodulin molecules from the myosin heavy chain or because of Ca^{2+}-induced allosteric effects in the calmodulin light chains (Wolenski 1995).

Following the neck domain, each myosin has some form of a class-specific tail domain. The tail domain is the most divergent part of the myosin molecule. It consists of coiled-coil α-Helix which promotes the formation of dimers and bipolar filaments (except the Myosin I monomer, which does not have a-helical repeats, but has a highly basic domain shown to bind to phospholipid membranes (Mooseker and Cheney 1995; Hayden et al. 1990). The function of the C terminal domain is associated with targeting of specific molecules to specific subcellular localisation (Mermall et al. 1998; Oliver et al. 1999).

3.1.2 Class V myosins

3.1.2.1 Domain structure of myosin V

The myosins from class V are two headed and do not appear to form bipolar filaments (Fig.3.1.2.1). The head, located at the N-terminus, contains the ATP hydrolysis site and the actin binding site. The neck domain consists of six IQ domains (12 in a dimer molecule) binding eight molecules of calmodulin and four light chain molecules of dynein or tissue specific myosin light chains (Espindola et al. 2000). The myosin V tail is composed of a coiled-coil domain, which allows dimer formation with a second myosin molecule, and a unique globular domain (Mooseker and Cheney 1995; Larson 1996). The sequence in the medial and distal parts of the tail domain are highly conserved among vertebrate myosins but show little similarity to myosins V tails from more phylogenetically diverse species (Fig. 3.1.2.1). The function of the globular domain is associated with directing myosin V to its targets and/or determining the cargo molecule to which it binds.

3.1.2.2 Distribution of myosin V

Class V myosins are one of the most well characterised groups of the unconventional myosins (Mooseker and Cheney 1995; Larson 1996; Titus 1997). Members have been cloned from the slime mould *D. discoideum* (*MyoJ*), yeast

(*Myo2* and *Myo4*), *C. elegans* (*hum-2*), frog (*myoV*), chicken (*p190*), mouse (*dilute*), rat (*myr6*) and human (*Myoxin*). Several plant myosins were also identified as class V myosins. These currently are placed in a separate class since their globular tails showed a number of differences when compared to other class V myosins. The vertebrates have three distinct myosin V genes, Va, Vb and Vc. Currently from the invertebrates only *C. elegans* is known to have two myosin V genes. Interestingly plants expresses a multiple of myosin V genes. So far have been identified 13 myosin V genes in *A. thaliana* (Berg et al. 2001).

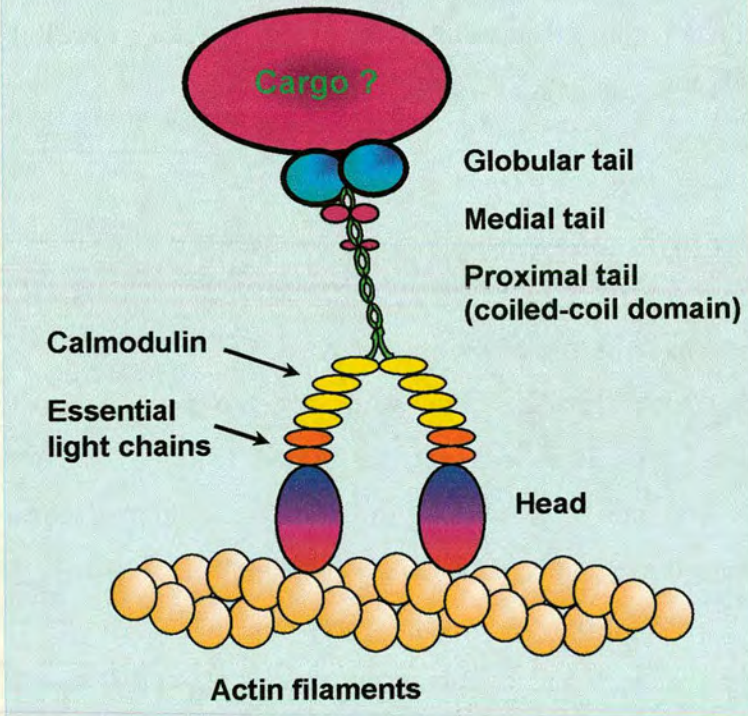


Figure 3.1.2.1 Structural 3D model of Myosin V

The myosin motor head domain associates with actin filaments. The neck binds eight molecules of calmodulin and four subunits of essential light chains. The globular tail domain is attached to its cargo.

3.1.2.3 Regulation of the motor activity of myosin V

Three main factors are necessary for the activation of the myosin molecule. As in all myosins the Mg-ATPase activity is induced by the addition of actin filaments. Ca^{2+} ions are also required for the Mg-ATPase activity and the activity is further stimulated by the presence of calmodulin (Wang et al. 2000)

At low Ca^{2+} levels the Mg-ATPase activity is also low. It was found that at high Ca^{2+} levels inhibition of the motility occurs too. Recent studies have shown that this is due to conformational changes of the calmodulin upon Ca^{2+} binding to the high affinity site, but not due to dissociation of calmodulin from the myosin heavy chain (Homma et al. 2000). Interestingly when Ca^{2+} is added to motility assays *in vitro*, the movement of the myosin molecules is inhibited (Cheney et al. 1993b; Wolenski et al. 1995)

3.1.2.4 Myosin V stepping kinetics

Myosin V moves processively along actin filaments. Kinetic studies with truncated molecules have greatly contributed to our understanding of the mechanism underlying ATP turnover and movement (De La Cruz et al. 2000). To understand the processive mechanism of myosin V Rief et al. (2000) has studied a full-length dimer molecule. He proposed a hand-over-hand stepping model where one hand is making a stroke while the other is reaching out for the next site, without loosing contact with the underlying actin network. Because of the long neck domains the heads of the two heavy chains in a dimer are spread apart from each other, thus allowing the molecule to reach adjacent actin filament. It is believed that this can facilitate cargo transport within a loose actin network.

3.1.3 Myosin V distribution and function

The function of any myosin is determined to a great extent by the composition of its tail domain. While the proximal coiled-coil domain has been implicated in the dimerisation of the myosin molecule the medial and distal globular domains are thought to determine the exact roles and subcellular locations of the myosin molecule. At present evidence suggests a role for myosin V in organelle and particle transport and RNA localisation (Evans et al. 1997; Mermall et al. 1998).

3.1.3.1 Myosin V genes in plants

Plant cells exhibit an active intracellular motility, based to a great extent upon actin-based processes, such as cytoplasmic streaming and directed vesicular transport (Kinkema et al. 1994). The motor proteins responsible for the actin-based transport are myosins. Plants express at least eight distinct classes of myosin genes. The first

member of the class V myosin was isolated from *Arabidobsis thaliana* (Kinkema and Schiefelbein 1994; Reddy and Day 2001). It is a 173 kDa protein structurally similar to class V myosins from animal species. Subsequently a myosin V was cloned from *Zea mays* (Liu et al. 2001), *Helianthus annuus* and *Chara corallina* (Kashiyama et al. 2000). Currently all class V myosins from plants are placed in a separate class, XI.

3.1.3.2 Myosin V genes in social amoebae

Cellular slime moulds are simple eukaryotic organisms. However they express multiple myosin genes (Soldati et al. 1999). A single myosin V gene, named *MyoJ* was isolated from *Dictyostelium* cells (Hammer and Jung 1996). It encodes a big 258 kDa protein which shows the typical domain structure for myosin V. MyoJ differs slightly from other myosin Vs since the head domain contains a 30 amino acid insert in the proximity of the nucleotide binding pocket. In addition the globular tail shows no similarity to tails from vertebrate myosin Vs. Phylogenetic analysis of the head domain has shown that MyoJ is more closely related to plant myosin Vs and currently is placed into class, XI.

3.1.3.3 Myosin V genes in yeast

Cells in *S. cerevisiae* divide by budding. The daughter cells differ only in that the mother cells express the *HO* endonuclease and thus switch mating types. Two Myosin V are present in the budding yeast, *Saccharomyces cerevisiae*. Myo2p is an essential myosin of fundamental importance for polarised growth. Cells lacking this myosin fail to deposit the chitin ring which is required for bud formation (Johnston et al. 1991; Brockerhoff et al. 1994). The cells remain as large unbudded cells showing accumulation of numerous vesicles. Temperature sensitive mutants are able to complete the cell division when shifted to restrictive temperatures after chitin ring formation, suggesting no role for this gene in nuclear migration and cytokinesis. Interestingly the Myo2p IQ motifs bind not only calmodulin but a novel protein termed Mlc1p from the EF-hand family (Stevens and Davis 1998). Thus IQ-associated chains are believed to regulate the motor activity.

MYO4 is a second class V myosin gene in yeast. Evolutionarily it is most close to *MYO2* (Haarer et al. 1994). It encodes a 170kDa protein with a short potential coiled-coil domain. Myo4p is necessary for restricted localisation of *Ash* mRNA in the daughter cell. As a result ASHIp, a repressor of mating type switching (repressor of *HO*), is selectively expressed in the daughter cells (Jansen et al. 1996). Deletions in *MYO4* yield no phenotype indicating that the function of the gene is not essential in yeast (Haarer et al. 1994; Mermall et al. 1998). However overexpression of either gene *MYO4* or *MYO2* results in a severe decrease in growth rate, as well as aberrant phenotypes such as formation of short strings of unseparated cells in diploid strains or clusters of cells in haploid strains.

3.1.3.4 Myosin V in nematodes

A single myosin V gene was isolated from the round worm *Caenorhabditis elegans* (Baker and Titus 1997). The gene *hum-2* codes for approximately 226 kDa protein. HUM-2 is closely related to vertebrate myosin V proteins. It has the typical head, six IQ domains and a tail with 3 coiled-coil motifs. The final part of the tail shows high homology to the tail domain of class V unconventional myosins. Since this is the domain responsible for the neurological function of the protein it is suggested that mutants with uncoordinated movements or a lethal phenotype are candidates for HUM-2 function (Baker and Titus 1997; Mooseker and Cheney 1995).

3.1.3.5 Myosin V in echinoderms

Sea urchin has a single myosin V gene (Sirotkin et al. 2000). It encodes a 210 kDa protein that shares a significant degree of sequence identity with studied myosin V proteins. Protein comparison of the tail domain with sequences from vertebrate myosin V showed 72% homology, which suggest that these domains may have similar function.

3.1.3.6 Myosin V in cephalopods

The first myosin V from cephalopods was cloned from the squid, *Loligo pealei* (Molyneaux and Langford 1997; Molyneaux et al. 2000). The protein isolated from a brain fraction is approximately 196 kDa (Cohen 2001) and is abundantly expressed in the squid giant axon. P196 has been implicated in the transport of

axoplasmic organelles along the actin filament in the axon. Molyneaux et al. (2000) has demonstrated that antibodies to myosin V inhibit this transport. Further research showed that purified p196 is a calmodulin (CaM) binding protein that possesses calcium-stimulated actin-activated ATPase activity.

3.1.3.7 Myosin V genes in amphibians

The proper function of the mechanoelectrical transduction mechanism in the hair cells of the auditory and vestibular systems in frogs appears to be dependent on the tension of the mechanosensitive transduction channels. Evidence suggests that members of the myosin family can control this tension. In a screen for myosin candidates for this function the frog homologue of myosin V gene was isolated (Solc et al. 1994).

It has been known that melanosome movements in amphibians and fish are dependant on microtubules and their associated motors. Rodionov et al. (1998) demonstrated that both actin and microtubule systems are involved in transport from the perinuclear region to a uniform distribution in the cytoplasm. In addition Rogers et al. (1999) has shown that the association of myosin V with melanosomes is cell cycle dependent. They found that metaphase extracts of myosin V are highly phosphorylated compared to interphase extracts. Thus the organelle transport by myosin V is controled by phosphorylation during mitosis.

3.1.3.8 Myosin V genes in birds

Low levels of myosin V are expressed in most vertebrate tissues, but it is most abundant in nervous tissue and neurosecretory cells (Espindola et al. 1992). Myosin V was identified in chicken as a 190 kDa protein (p190) (Espreafico et al. 1992). Immunodetection of p190 in brain showed that the protein is most abundant in the cell bodies and dendrites of Purkinje cells. In tissue cultures of neurons and glia, myosin V shows a characteristic distribution in the region of the Golgi complex and in the tips of the neuronal growth cones (Espreafico et al. 1992). Purification of chicken Myosin V allowed electron microscopic visualisation of the two headed dimer it forms. It was also found that chick myosin V, when bound to actin, exhibits a magnesium-ATPase activity (Cheney et al. 1993b).

This myosin was demonstrated to have the unusual property of binding calmodulin in the absence of calcium (Cheney et al. 1993a; Nascimento et al. 1997). Analysis of purified chick brain myosin V revealed that four of the IQ domains bind calmodulin. Two new proteins, subunits of 17 and 23 kDa, have been found to associate with the remaining IQ motifs. The 17 and 23 kDa proteins are encoded by *L17* and *L23* genes and are members of the essential light chain family. In myosin II they bind to the first of the two IQ domains. An additional 10 kDa protein subunit was found to co-immunoprecipitate with the globular tail domain. This was first characterised as DLC8, one of the axonemal and cytoplasmic dynein light chains. Subsequently it was found that DLC is a potential inhibitor of the neuronal form of NO synthase (PIN). It was demonstrated by Crepieux et al. (1997), that DLC8/PIN interacts with I $\kappa B\alpha$ transcriptional regulator as well as with Bim, a member of the Bcl-2 family of membrane proteins. These additional proteins are believed to play critical role in the structure and function of myosin V (Espindola et al. 2000).

The role of Myosin V in growth cone dynamics was examined by a laser inactivation technique (Wang et al. 1996). Laser inactivation of Myosin V inhibited the ability of myosin-V to translocate actin filaments resulting in rapid filopodial retraction, suggesting a specific role of this myosin in filopodial extension.

3.1.3.9 Myosin V genes in mammals

The mouse *dilute* gene was the first member of the family to be cloned (Mercer et al. 1991; Espreafico et al. 1998). In normal cells the melanocytes synthesise pigment packed in melanosomes. Then they deliver the melanosomes by means of dendritic processes to the keratinocytes. Mutations in *dilute* lead to defective pigment granule (melanosomes) transfer which results in lightening of the coat colour from black to grey. Recent studies revealed that melanin is produced at normal levels but accumulates in the perinuclear region of melanocytes. Null mutations of *dilute* cause severe neurological disorders (Provance et al. 1996; Wu et al. 1997). After birth these mice exhibit ataxia, abnormal behaviour, disrupted balance and convulsions which lead to early postnatal death.

A second dilute-like myosin, *myr 6* has been isolated and characterised from rat (Zhao et al. 1996). A homologue of *myr 6* has been found in mouse (McIver and

Bownes, unpublished). Since the sequence of *myr 6* is very similar to that of dilute it is believed that the two genes will have overlapping functions. Consistent with this model is the presence of DIL (AF-6/canoe) domain in the tails of both genes. AF-6 and canoe are PDZ/DHR (Disc-large homology region/PSD95, Disc-large, ZO-1) domain proteins participating in cellular junction formation, receptors or channel clustering and signal transduction.

Myoxin (MYH12), the human homologue of the mouse *dilute* gene was identified by Engle and Kennett (1994) and Moore et al. (1995). The gene was mapped to chromosome 15. Recent studies have shown that in human, as in mouse, there are a number of tissue specific alternative-splice forms of the myosin V gene. All transcripts (different combinations of A-B-C-D-E-F exons) are present in skin cells, particularly in the melanocytes, while brain transcripts lack exon F. This produces a myosin V with a truncated C-terminal tail. Since this is the region where the synaptic vesicles bind to myosin V (by means of the membrane proteins synaptobrevin and synaptophysin) it was suggested that the binding to different targets is regulated by producing different splice variants (Prekeris and Terrian 1997). Humans with Griscelli syndrome have pigmentary dilution, in which the melanosomes are not present in the melanocytes dendrites, but are clustered in the perinuclear region. Linkage analysis has placed Griscelli syndrome in the proximity of myosin V. Two thirds of the patients with Griscelli syndrome are found to have mutations in the gene for myosin V.

3.2 *Drosophila Myosin V*

A number of enhancer trap lines containing P[*lacZ;rosy*] insertions have been screened in our laboratory. Several lines with staining patterns of interest in the ovaries were selected for further study. One of these lines demonstrated a temporal and spatial expression pattern in a subset of follicle cells. The target gene was cloned but its expression pattern was unrelated to the original reporter gene expression pattern (MacIver et al. 1998). However the revealed pattern was highly interesting. *In situ* hybridisation with the rescued flanking DNA showed that the transcripts were localised to the anterior region of the oocyte from very early in oogenesis (stages 3 to 7). Later stages (9-11) showed strong expression in the nurse cells, and a strong

anterior band in the oocyte. The pattern in the oocyte is similar to that seen for the *orb* and *oskar* mRNAs suggesting a function in oogenesis and/or embryogenesis. The correct transcript localisation was found to require genes involved in oocyte determination, anterior and posterior patterning, as well as neurogenic genes involved in oocyte-follicle cells signalling. Molecular analysis showed that the newly cloned gene, named *didum*, encodes a class V unconventional myosin (MacIver et al. 1998; McCormack et al. 1998).

A developmental profile of *didum* by Northern blotting showed strong expression in adult ovaries and testes with transcripts of 5.5kb and 6kb respectively. Low levels of expression were detected throughout the development in non-gonadal tissues (McCormack, personal communications)..

The gene was mapped to position 43Cc on the right arm of the second chromosome. All known genes corresponding to position 43Cc (position of saturation mutagenesis, are essential genes and lead to an embryonic lethal phenotype (Heitzler et al. 1993).

d*idum* encodes a 1792-residue, 207kDa heavy chain polypeptide; an actin-binding head domain of 771-residue, a neck region with six IQ motifs (139-residue), and a tail domain of 882-residue with coiled-coil motif required for dimerisation (MacIver et al. 1998; Bonafe and Sellers 1998). A unique substitution for *Drosophila* Myosin V is the replacement of the third IQ motif with LS (leucine, serine) (MacIver et al. 1998). The domain structure of *Drosophila* myosin V will be discussed in detail in chapter five.

A region from the C-terminal tail of the gene was expressed in *Escherichia coli* and the recombinant protein was used to raise antibodies in rabbits (MacIver et al. 1998). Immunolocalisation with the unpurified antibodies showed no specific staining, but diffuse overall cytoplasmic staining. This together with the fact that extensive washing with detergent removes most of the staining, suggests the antibodies bind unspecifically.

From the loss of transcript localisation in ovaries treated with colchicine it could be concluded that an intact microtubule cytoskeleton is required for localisation of myosin V mRNA (McCormack, personal communication).

3.3 Determining the chromosome position of didum

The *Drosophila* myosin V gene was previously mapped to position 43Cc on the right arm of the second chromosome (Bonafe and Sellers 1998; MacIver et al. 1998). It was named *didum* for *dilute*-like *Drosophila* unconventional myosin. With the release of the *Drosophila* Genome Sequence we were able to precisely map the gene to position 43D1. *didum* is positioned immediately downstream (genetically) of *dpa (disc proliferation abnormal)* with a known chromosome position, 43D1. We verified its exact location by genomic PCR (the template is genomic DNA) using specific for the *dpa* forward primers (DP1 and DP2) and a reverse primer for *didum,* RACE3 (for primer sequences and information see Fig. 3.3B). As a control a primer from the 5' end of *didum* (5'UTR) combined with the same RACE3 primer (Fig.3.3A) was used.

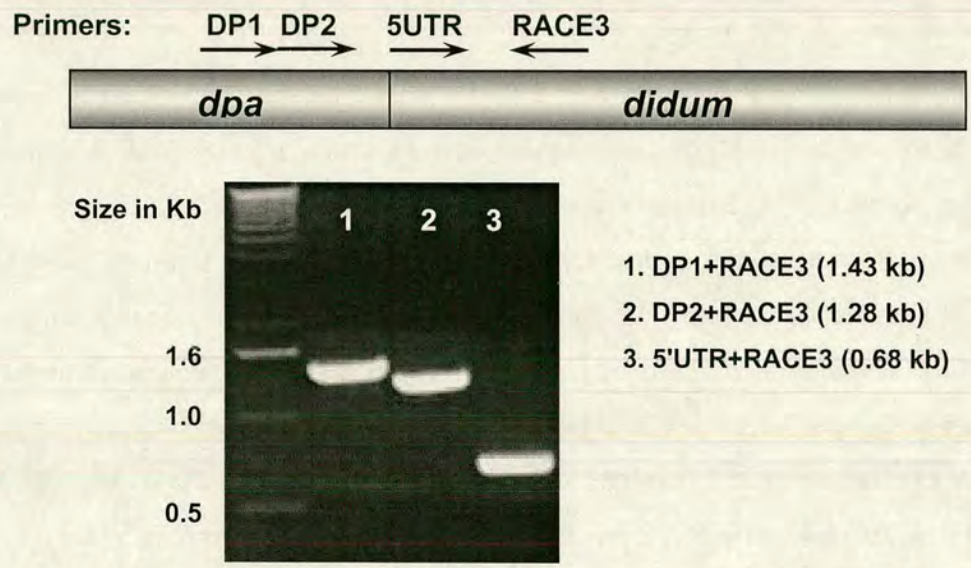


Figure 3.3A Genomic PCR for verifying the position of *didum*

Three different primer combinations were use to verify the genomic position of *didum*. The primers used in the PCR are depicted by arrows. The products amplified by the PCR were with the expected sizes thus confirming our predictions. The PCR was carried out as follows: 94°C-1.5 min, then 25 cycles at 94 °C - 30 sec., 55 °C - 30 sec. and 70 °C - 4 min. and a final extension for 4 min. at 70 °C. The PCR amplified single products with the expected size of 1.43 kb (for the DP1+ RACE3), 1.28 kb (DP2+RACE3) and 0.68 kb for the control (5'UTR+RACE3). The obtained products were excised from gel, purified and sequenced. The sequencing results confirmed the position of *didum* in relation to *dpa*.

Figure 3.3B Primers used in the analysis of *didum* (Myosin V):

Primer name	Sequence in 5' > 3' direction	Length (nt)	Tm (°C)
DP1	GTC TCC ACA TCT TGC GAC G	19	60
DP2	GGC TTG ACA TTT CCT TAG AGA	21	60
5UTR	AAC CTA ACG GGA TTT ATG CAC	21	60
BSM1	CCG GCT TCT TGA AGA TAT G	19	56
STM1	GGC GTA CAC AAA ACT GGA G	19	58
STM2	AAT ATG TGA TAG TTG CGG TCT C	22	62
RACE3	GCC TTT AGC TTG ACC TCT TTC	21	62
did8	ACA ACA GTT CCC GCT TTG	18	54
did5	CCA TTC AAC AGA TCG CTG	18	54
SH-1R	CGA GGG AGT TGG CGG TGA	18	60
SH-2R	GGC GAA ATT AGT AGG TCG TG	20	60
did9	ACC AGC ATG TCT TCA AGC	18	54
did10	TCC AAG AAT CCG TTC ACG	18	54
did4	TTC CGC TAA ATC CTC CAC	18	54
did3	AGT TTG GAG CGG TAA ACC	18	54
1806	GGC CAC GTC GAT TGG AGC	18	60
926Y	GCG GAT TAT CAT GCC AGT C	19	58
MR3	GGC TTG CAT AGT CTA ACG AC	20	60
3'UTR5	TTT AAA CGT CGC GTG CCT	18	54
3'UTR3	TAA TAT ATA ATA TTT GTA TGT	21	46

Using the *Drosophila* Genome sequence we were able to locate *myosin V* gene (*didum*) on the (+) strand of DNA. The position of *didum* in relation to the nearby genes is shown on Fig. 3.3C.

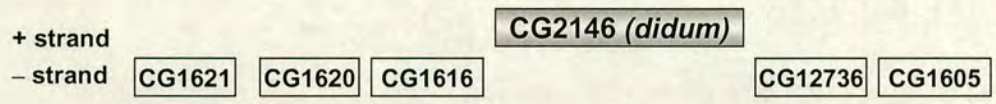


Figure 3.3C. Genomic position of *didum*

The diagram shows the relative position of *didum* in the region 43CD on the right arm of chromosome 2. CG1621 is a transcription factor, CG1620 is a replication factor, CG1616 (DPA) regulates initiation of DNA synthesis, CG12736 is a translation factor, and CG 1605, also known as *az2*, is a zing finger transcription factor (McCormack et al. 1998).

3.4 *Molecular structure of didum*

Most of the myosin V sequence including the ORF (open reading frame) has been cloned and published (MacIver et al. 1998) and (Bonafe and Sellers 1998). Previous attempts to sequence the very 5' end of *didum* were not successful. From a sequence analysis of the products generated by PCR for verifying the position of *didum* we were able to determine the potential start site of the gene. We used the Neural Network Promoter Prediction Program, available at the page of the Berkeley *Drosophila* Genome Project, (http://www.fruitfly.org/seq_tools/promoter.html) to find the possible promoter sites. Setting the default parameters produced several high score results which were inconclusive because the predicted promoter sites were obviously from the ORF (open reading frame) sequence. When the score cut off limit was lowered we obtained a number of low value scores in the expected region (the lower the score cutoff, the more potential promoters will be shown). To verify their accuracy we analysed the promoters of several *Drosophila* genes (with known promoters) and compared them to the predicted *didum* promoter sequence (Fig. 3.4A).

Gene	Score	Promoter sequence	Transcription start
Dm EF-1a'F1	0.89	AAAAAATG**TATAA**ATAAATGAAACGTATTACATAAATTTA**A**ACGATCTCA	
Dm EF-1a'F2	0.44	TTTAAGAGAA**TATAA**AATATTCCATGAATGGTAGTAAAATT**T**GTATTACTA	
Dm[42A]actin cp	1.00	TAAAAAATTT**TAAAAA**AAGTCCGCTCTCCAGTCTTCACCG**T**TTCCAACTT	
Dm[5C]actin E1+	1.00	AAGCGGGCTT**TATAA**AACGGGCTGCGGGACCAGTTTTCAT**A**TCNNNNNNN	
Dm[5C]actin E2	0.77	CGAAACATGACCA**TATA**TGGTCACAAAACTTGGCCGCCG**CA**ATTCAACAC	
Dm[87E]actin(muscl)	0.99	CAGAATGG**TATAA**ATATTAGCGCATCTCGGTCCAGCGAC**CA**CTCGCAGTT	
Dm tropomyosin I	0.50	AGAAAAGCG**CATAAA**TTTCACGAGAATGCAGCCCAAAAGAGAGAGCATCG	
Dm paramyosin P1	0.67	CTCAGGG**TATA**TCCCACTCGGATCCGATCGGCGGGGAGC**CG**ATGGCGCCC	
Dm paramyosin P2	0.67	CTCAGGG**TATA**TCCCACTCGGATCCGATCGGCGGGGAGC**CG**ATGGCGCCC	
Dm miniparamyosin	0.92	GTGCGACAA**TATAA**CGCCGACGAGATCGTCCGATTGTCG**CA**CCTGCGCCG	
Dm chorion p.s38	0.86	TCCGGAG**CATA**TTTAAAGTAGTCGGCCACCAATGGAGGG**CA**GCAGAAGAC	
Dm chaoptin	0.91	CCATATTC**TATA**CAAAGGCCAGCATGACCAGCATTTGGC**CA**TTCCTCCAG	
Dm calmodulin	0.97	AATAAAACCG**TATA**TATGGCGACGCAGCGCTATTTTGAA**CA**CAAAGTTTT	
Dm[Krueppel]	0.82	NGTAGACTTT**TATAA**AAGACAATTTTTGTGAAATCTCTCT**A**CCTCAAAGT	
Dm[Krueppel2]	0.27	ACCTCAAAG**TACAA**AAGTGTGTACAAAAATTATTCATAT**CC**CTGAAAGTG	
Dm[engrailed]	0.73	ATATTGCTGCC**TAACAA**GCCCCCCTCGCCAAATTGGCAAG**A**GAGAGTGAG	
Dm[ninaE]opsin/Rh1	0.99	CATTGCGAAAAC**TATAA**AAGCCACGCGGCAGAATGCAGA**CA**TTGCAGGTT	
Dm Myosin V (I)	0.18	TGCCAACAACAAAA**CATAA**GTCAGTCAGTCAGGCGGTGA**CA**GGGACGGAA	
Dm Myosin V (II)	0.48	AAATTATCAATGTTTT**TATAAA**TTATTCCTGGCTAATTTT**A**AGTAATTTA	
Dm Myosin V (III)	0.81	CATTTA**CATAAA**CCCAAAAAAAGACCACGAAAATGGTAATT**T**AAATAGGC	
Dm Myosin V (IV)	0.47	ACCACGAAAATGG**TAATTTA**AATAGGCTGTCAGAGCCAGA**A**AAAAAACGT	

Figure 3.4A Alignment and comparison of selected *Drosophila* promoters

A number of genes were selected from the *Drosophila* promoter database (http://www.fruitfly.org/seq_tools/datasets/Drosophila/promoter/) and compared. The program is based on an algorithm recognising functional sites in the primary DNA that are involved in the polymerase binding, such as GC box, TATA box, CAAT box and transcription start. The data was used to clarify the possible promoter start sites of *didum*. The TATA binding sites are shaded in blue, the transcription starts sites are coloured in red.

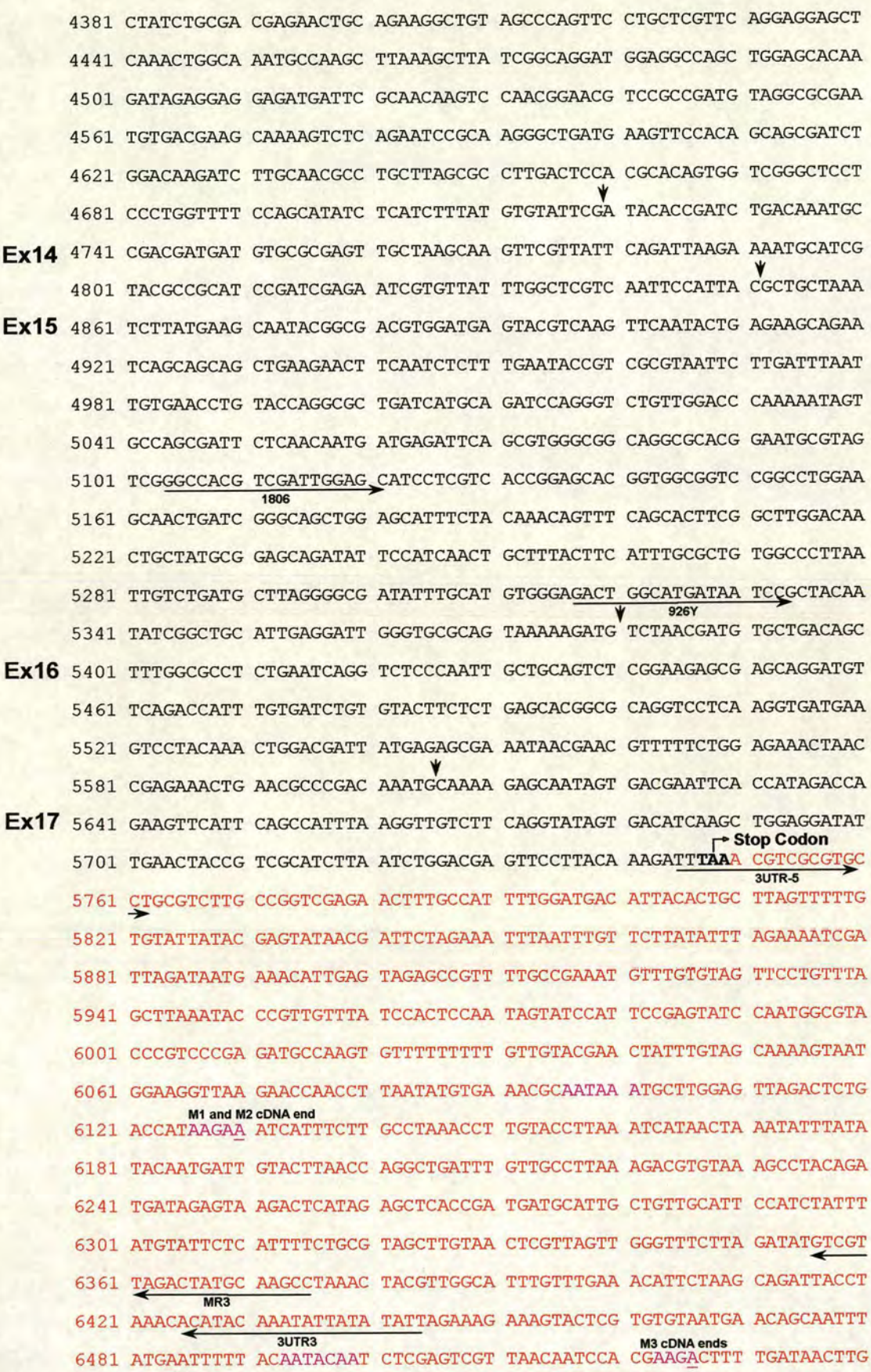
Four potential promoters were found with scores of 0.18, 0.48, 0.81, and 0.47 respectively (Fig. 3.4B). The Ps (promoter score) 0.18 and Ps 0.43 are the best candidates because they show comparatively high similarity to the conserved promoter sequence, and coincide with the experimental data from the isolated cDNA clones (for more details see next subchapter).

Figure 3.4B cDNA (mRNA) sequence for *didum*

Translated sequence is shown in black. Untranslated sequences (5'UTR and 3'UTR) are coloured in red. The promoter sequences are given in blue and the cut off scores are shown above the promoter sites. The transcription start sites (there are two transcriptions sites) are underlined and coloured in blue. The polyadenylation signals (AATAAA) and the subsequent cleavage sites (AAGA) are coloured in purple. The PCR primers used in the analysis of the gene are shown with arrows. →

Figure 3.4B cDNA (mRNA) sequence for *didum*


2161 CAAGCAGGTG ATGACCCTGG AGGAAATAGA TACTCTGTGC GTGGATTCCG CTAAATCCTC
did 4
2221 CACCTTGGGC GGCCGCGTCG TGATCAGTGC TGGCCGCAAA CAGGTGGTGC CATCCAAGCA
Ex7 2281 GCATAGAAAA ACGGTGGGAT CGCAGTTCCA GGAGAGTCTG GCGTCGCTGA TATCTACGTT
2341 ACATGCCACA ACTCCGCACT ATGTGCGCTG CATCAAGCCC AACGATGACA AAGTCGCCTT
Ex8 2401 TAAGTGGGAG ACGGCCAAGA TCATACAGCA GTTAAGGGCC TGTGGTGTGC TGGAAACGGT
2461 GCGCATCTCC GCAGCGGGAT TCCCCTCGAG ATGGCTCTAT CCCGACTTCT ATATGCGGTA
2521 CCAGCTGCTG GTTTACCGCT CCAAACTCGA CAAAAACGAC ATGAAGCTGT CGTGCCGGAA
did 3
2581 CATTGTGATG AAGTGGATCC AAGACGAAGA TAAGTACCGA TTTGGCAACA CGCAGATTTT
2641 CTTCCGCGCC GGCCAAGTGG CCTTCCTTGA ACAGGTTCGG GCTAATCTGC GCAAGAAGTA
2701 TATCACCATT GTGCAGTCGG TTGTGCGGCG ATTCGTCTAC CGGCGCCAGT TCCTGCGCAT
2761 TCAGAAAGTA ATTAATGGCA TTCAGAAACA TGCGCGCGGA TATCTTGCTC GCGAGCGTAC
did 11
2821 TCAGAAAATG CGCGAAGCTC GTGCGGGATT AATCCTGTCG AAGTACGCCC GCGGTTGGTT
2881 ATGCCGTCGT CGTTACTTGC GCCTACGCCA CTCCATTTCC GGCATACAGA CCTACGCCCG
2941 CGGCATGCTG GCGCGCAACA AGTTCCACGC GATGCGGGAT CACTACCGGG CAGTTCAGAT
3001 CCAGCGTTTC GTGCGTGGTG CTTTGGCACG GCGAGCTTAC CAAAAGCGGC GTCGCAACAT
3061 CATCATTTGT CAAGCGGCGA TTCGGAGATT CTTGGCCCGT CGTAAGTTTA AACGCATGAA
3121 GGCCGAGGCC AAGACCATCT CGCACATGGA AAACAAATAC ATGGGGCTGG AAAACAAGAT
3181 TATTTCCATG CAGCAGCGGA TCGATGAGCT GAATCGCGAC AACAGTAATC TGAAGCACAA
3241 GACCAGCGAA ATCAGTGTAT TGAAAATGAA GCTTGAGCTG AAGAAGACCC TGGAGGCTGA
Ex9 3301 ATTCAAAAAT GTCAAGGCCG CCTGCCAGGA CAAGGACAAG CTGATCGAAG CACTTAACAA
3361 GCAGTTGGAG GCGGAGCGAG ACGAAAAAAT GCAGTTGCTG GAGGAGAACG GACATGCTCA
3421 AGAGGAGTGG ATCAGCCAGA AGCAGACGTG GCGCCAAGAG AACGAGGAGC TGCGCCGTCA
3481 GATAGACGAG ATAATCGATA TGGCAAAGAA CGCAGAAGTC AACCAGCGTA ACCAGGAGGA
3541 CCGAATGCTA GCCGAGATTG ATAACAGGGA GCTCAACGAG GCCTACCAAC GAGCTATTAA
3601 GGACAAGGAG GTCATCGAGA ACGAAAACTT CATGCTGAAG GAAGAGCTCA GTCGATTAAC
3661 GGCTGGCAGT TTCAGTTTGC ACGGCCGCAA GGCTAGCAAC GCCTCCAGCC AAAACGAGGA
3721 CGATGTGGGA TACGCCTCCG CCAAGAACAC TCTGGATATC AATCGGCCCC CGGATTTGTT
3781 AAGCAAAAAT TACTCGTACA ATGACTCTAC CAGTCTGGTG GTGAAGTTGA GATCCATTCT
Ex10 3841 CGAGGAGGAG AAGCAAAAGC ACAAGGTCTT GCAGGAGCAG TACATTAAGT TGTCCAGTCG
3901 GCATAAGCCC ACCGAGGATT CCTTCCGCGT CTCCGAGCTT GAGGTAGAGA ATGAAAAGCT
Ex11 3961 GCGCAGCGAG TACGATCAGC TGCGAACGAG CATTAAACAC GGTGTTGAGA TCAACGAGCT
4021 CAATGCACAG CATGCCGCCT TGCAGGAAGA GGTACGTAGG CGGCGCGAGG AGTGCATCCA
Ex12 4081 ATTAAAGGCA GTCCTGCTGC AGCAGAGCCA GTCCATGAGA TCGCTCGAGC CGGAAAGTCT
4141 ACAGATGCGT GGCAACGATG TCAACGAACT GATGGAAGCC TTCCATTCCC AGAAGCTAAT
4201 TAATCGTCAA TTGGAGTCTG AGCTCAAGGC CATCACCGAG GAGCACAACA GTAAGCTCGT
EX13 4261 GGAGATGACA CAGGAGATCG AGAGATTGAA CAATGAGAAG GATGAGCTGC AAAAGGTAAT
4321 GTTCGAGAGC ATCGACGAGT TCGAAGATTC CAATGTGGAT ACGCTGAGAC AGAACGATCG


4381 CTATCTGCGA CGAGAACTGC AGAAGGCTGT AGCCCAGTTC CTGCTCGTTC AGGAGGAGCT
4441 CAAACTGGCA AATGCCAAGC TTAAAGCTTA TCGGCAGGAT GGAGGCCAGC TGGAGCACAA
4501 GATAGAGGAG GAGATGATTC GCAACAAGTC CAACGGAACG TCCGCCGATG TAGGCGCGAA
4561 TGTGACGAAG CAAAAGTCTC AGAATCCGCA AGGGCTGATG AAGTTCCACA GCAGCGATCT
4621 GGACAAGATC TTGCAACGCC TGCTTAGCGC CTTGACTCCA CGCACAGTGG TCGGGCTCCT
4681 CCCTGGTTTT CCAGCATATC TCATCTTTAT GTGTATTCGA TACACCGATC TGACAAATGC
Ex14 4741 CGACGATGAT GTGCGCGAGT TGCTAAGCAA GTTCGTTATT CAGATTAAGA AAATGCATCG
4801 TACGCCGCAT CCGATCGAGA ATCGTGTTAT TTGGCTCGTC AATTCCATTA CGCTGCTAAA
Ex15 4861 TCTTATGAAG CAATACGGCG ACGTGGATGA GTACGTCAAG TTCAATACTG AGAAGCAGAA
4921 TCAGCAGCAG CTGAAGAACT TCAATCTCTT TGAATACCGT CGCGTAATTC TTGATTTAAT
4981 TGTGAACCTG TACCAGGCGC TGATCATGCA GATCCAGGGT CTGTTGGACC CAAAAATAGT
5041 GCCAGCGATT CTCAACAATG ATGAGATTCA GCGTGGGCGG CAGGCGCACG GAATGCGTAG
5101 TCGGGCCACG TCGATTGGAG CATCCTCGTC ACCGGAGCAC GGTGGCGGTC CGGCCTGGAA
1806
5161 GCAACTGATC GGGCAGCTGG AGCATTTCTA CAAACAGTTT CAGCACTTCG GCTTGGACAA
5221 CTGCTATGCG GAGCAGATAT TCCATCAACT GCTTTACTTC ATTTGCGCTG TGGCCCTTAA
5281 TTGTCTGATG CTTAGGGGCG ATATTTGCAT GTGGGAGACT GGCATGATAA TCCGCTACAA
926Y
5341 TATCGGCTGC ATTGAGGATT GGGTGCGCAG TAAAAAGATG TCTAACGATG TGCTGACAGC
Ex16 5401 TTTGGCGCCT CTGAATCAGG TCTCCCAATT GCTGCAGTCT CGGAAGAGCG AGCAGGATGT
5461 TCAGACCATT TGTGATCTGT GTACTTCTCT GAGCACGGCG CAGGTCCTCA AGGTGATGAA
5521 GTCCTACAAA CTGGACGATT ATGAGAGCGA AATAACGAAC GTTTTTCTGG AGAAACTAAC
5581 CGAGAAACTG AACGCCCGAC AAATGCAAAA GAGCAATAGT GACGAATTCA CCATAGACCA
Ex17 5641 GAAGTTCATT CAGCCATTTA AGGTTGTCTT CAGGTATAGT GACATCAAGC TGGAGGATAT
Stop Codon
5701 TGAACTACCG TCGCATCTTA ATCTGGACGA GTTCCTTACA AAGATTTAAA CGTCGCGTGC
3UTR-5
5761 CTGCGTCTTG CCGGTCGAGA ACTTTGCCAT TTTGGATGAC ATTACACTGC TTAGTTTTTG
5821 TGTATTATAC GAGTATAACG ATTCTAGAAA TTTAATTTGT TCTTATATTT AGAAAATCGA
5881 TTAGATAATG AAACATTGAG TAGAGCCGTT TTGCCGAAAT GTTTGTGTAG TTCCTGTTTA
5941 GCTTAAATAC CCGTTGTTTA TCCACTCCAA TAGTATCCAT TCCGAGTATC CAATGGCGTA
6001 CCCGTCCCGA GATGCCAAGT GTTTTTTTTT GTTGTACGAA CTATTTGTAG CAAAAGTAAT
6061 GGAAGGTTAA GAACCAACCT TAATATGTGA AACGCAATAA ATGCTTGGAG TTAGACTCTG
M1 and M2 cDNA end
6121 ACCATAAGAA ATCATTTCTT GCCTAAACCT TGTACCTTAA ATCATAACTA AATATTTATA
6181 TACAATGATT GTACTTAACC AGGCTGATTT GTTGCCTTAA AGACGTGTAA AGCCTACAGA
6241 TGATAGAGTA AGACTCATAG AGCTCACCGA TGATGCATTG CTGTTGCATT CCATCTATTT
6301 ATGTATTCTC ATTTTCTGCG TAGCTTGTAA CTCGTTAGTT GGGTTTCTTA GATATGTCGT
6361 TAGACTATGC AAGCCTAAAC TACGTTGGCA TTTGTTTGAA ACATTCTAAG CAGATTACCT
MR3
6421 AAACACATAC AAATATTATA TATTAGAAAG AAAGTACTCG TGTGTAATGA ACAGCAATTT
3UTR3
M3 cDNA ends
6481 ATGAATTTTT ACAATACAAT CTCGAGTCGT TAACAATCCA CGAAGACTTT TGATAACTTG

Analysis of the 3'UTR sequence showed two potential polyadenylation sites. These were confirmed when cDNA clones for the gene were isolated (Fig. 3.4B) (see next chapter for details).

3.5 *Myosin V transcripts and their localisation*

3.5.1 Isolation and analysis of *didum* cDNA clones

To study the molecular organisation of *didum* transcripts three new cDNA clones were isolated from an ovarian cDNA library (this was done by Timothy Wood). Initially the clones were analysed with PCR with specific primers for *didum*. The results (not shown here) indicated that these are full length clones based on the known sequence for the gene (Fig. 3.5.1A).

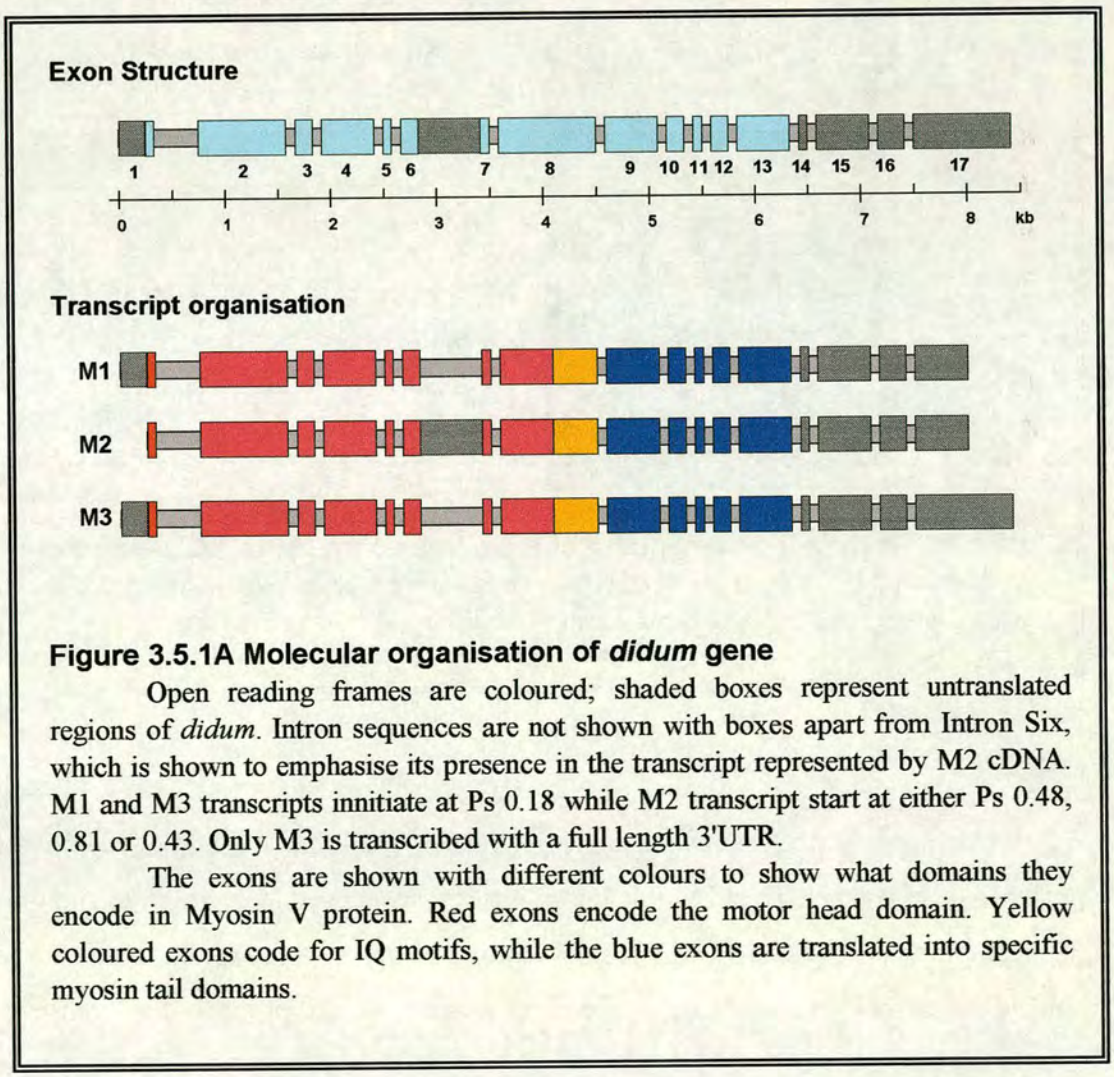


Figure 3.5.1A Molecular organisation of *didum* gene

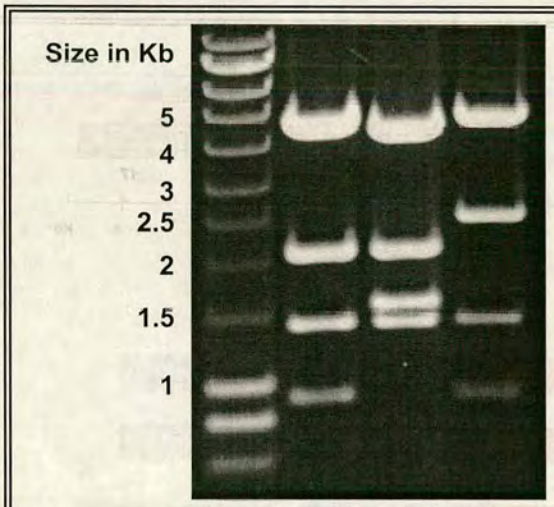
Open reading frames are coloured; shaded boxes represent untranslated regions of *didum*. Intron sequences are not shown with boxes apart from Intron Six, which is shown to emphasise its presence in the transcript represented by M2 cDNA. M1 and M3 transcripts innitiate at Ps 0.18 while M2 transcript start at either Ps 0.48, 0.81 or 0.43. Only M3 is transcribed with a full length 3'UTR.

The exons are shown with different colours to show what domains they encode in Myosin V protein. Red exons encode the motor head domain. Yellow coloured exons code for IQ motifs, while the blue exons are translated into specific myosin tail domains.

The 5' and 3' sequence of each of the clones were sequenced. The sequencing results showed that M1 and M3 clones ("M" for Myosin V) start at promoter Ps 0.18 while M2 utilises one of the other potential promoter sites, Ps 0.48, 0.81 or 0.43 (Fig. 3.4), most likely the Ps 0.43 since the start of the isolated cDNA matches exactly with this region (Fig3.5.1A)

The sequence at the 3' end showed that M1 and M2 have shorter 3'UTR which match well with the predicted two polyadenylation sites (Fig. 3.5.1A).

Restriction analysis of the three clones showed that the M2 transcript contains an internal intron. When translated this will shift the ORF (open reading frame) and introduce a STOP codon, thus producing a truncated form of myosin V protein (Fig. 3.5.1B).



XhoI digest, products (kb):

M1	M2	M3
4.49	4.23	4.49
2.02	2.02	2.38
1.36	1.5	1.36
0.88	1.36	0.88
0.29	0.29	0.29
		0.05

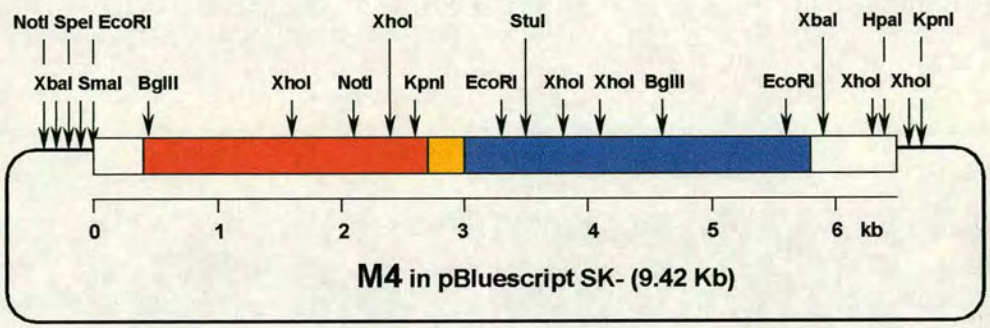
Figure 3.5.1B Restriction mapping of the obtained myosin V cDNA clones

Plasmid DNA from the three cDNA clones were digested with XhoI. M2 clone showed a larger band of ~1.5 kb compared to the expected 1.36 kb product.

Analysis of the *didum* sequence indicated that this might be due to the presence of Intron six. To check this possibility we sequenced M2 with did4 and did3 primers (shown on Fig3.4B). The results confirmed our hypothesis that Intron six is spliced together with the exons encoding myosin V to produce the transcript represented by the M2 clone.

3.5.2 Construction of M4 hybrid clone

M1 and M3 have identical start sides but the sequences were found to differ slightly at the 5' end at a site in the polylinker of the cloning vector. Thus, the M3 cDNA clone lacked the very 5'EcoRI restriction site. As this site was needed for cloning experiments (see chapter five) we created a hybrid between M1 and M3 called M4. M1 and M3 were digested with StuI and SpeI endonucleases and the fragments purified from agarose gel. The 5' part of M1 was ligated to the 3' part of M3 (the part containing the cloning vector). Thus the produced hybrid contain the 5' EcoRI restriction site and the full length 3'UTR belonging to M3 cDNA (Fig. 3.5.2).



Bgl II	5.20 + 4.22 kb		
EcoRI	3.86 + 3.24 + 2.32 kb	***SpeI***	9.4 kb
HpaI	9.4 kb	***StuI***	9.4 kb
KpnI	5.35 + 4.05 kb	***XbaI***	5.81 + 3.61 kb
NotI	7.22 + 2.20 kb	***XhoI***	4.47 + 2.37 + 1.36 + 0.88 + 0.29 + 0.05 kb

Figure 3.5.2 Restriction map of the hybrid M4 clone

The hybrid was constructed by combining the 5' end of M1 containing the *EcoRI* restriction site and exons coding for head and neck domains and the 3' end of M3, containing the exons for the tail domains, the 3'UTR and the vector sequence.

3.5.3 Northern analysis of *didum* transcripts

To find out the actual number and size of *didum* transcripts we used Northern analysis. 15μg of mRNA of different developmental stages were electrophoresed and transferred to a nylon membrane. The blotted mRNAs was hybridised with a probe

made to specifically detect the myosin tail sequence. The probe was generated by PCR using did11 and MR3 primers and radioactively labelled.

The analysis showed that all studied stages abundantly express myosin V. Two transcripts were detected: a 6.0 and 6.5 kb. The 6.5 kb mRNA is expressed at all developmental stages. The 6.0 kb transcript showed lower levels of expression in testes and male carcasses, and increased levels during larval stages (Fig. 3.5.3). As a loading control we used an Rp49 probe. Rp49 is a ribosomal protein expressed throughout the development at comparatively constant levels. We used PCR and specific primers (RP491 and RP492) to generate the control probe. These primers should produce a 400 bp PCR product. The PCR was carried out as follows: 94°C-1.5 min, then 30 cycles at 94 °C - 30 sec., 62 °C - 30 sec. and 70 °C - 1 min. and a final extension for 4 min. at 70 °C.

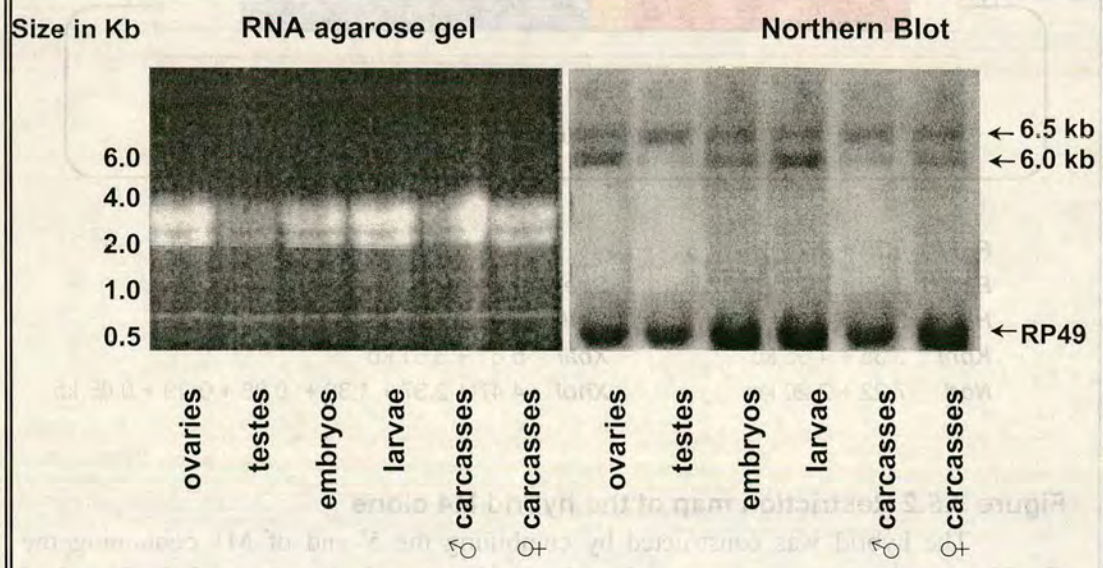


Figure 3.5.3 Developmental northern analysis of *didum* expression

The first picture shows the denaturing agarose gel used to separate the mRNAs. Total RNA from *D. melanogaster* should have two heavy ribosomal RNA bands at about 3.9 kb and 2.0 kb, and a lighter smear of mRNA centered around 2.0 kb (Tautz et al. 1988). In fact the 3.9 rRNA splits into two subunits, under denaturing conditions, to produce 1.7 kb and 2.2 kb products. The 2.2 kb and 2.0 kb bands partially overlap producing a wider band on agarose gels and a separate 1.7 kb band. Two myosin V transcripts were detected (6.0 kb and 6.5 kb) differentially expressed throughout the *Drosophila* development. RP49 was used as a loading control.

3.5.4 ESTs analysis

A BLAST search using the complete myosin V sequence was done to identify all EST sequences that match with *didum*. The database search showed that there are at least three different groups of ESTs available: sequences that contain all exons and UTR regions; ESTs with a short 3'UTR region and ESTs lacking exon 11. This data was combined with the information obtained from the analyses of the isolated *didum* cDNA clones to summarise all possible transcript variants (Fig. 3.5.4).

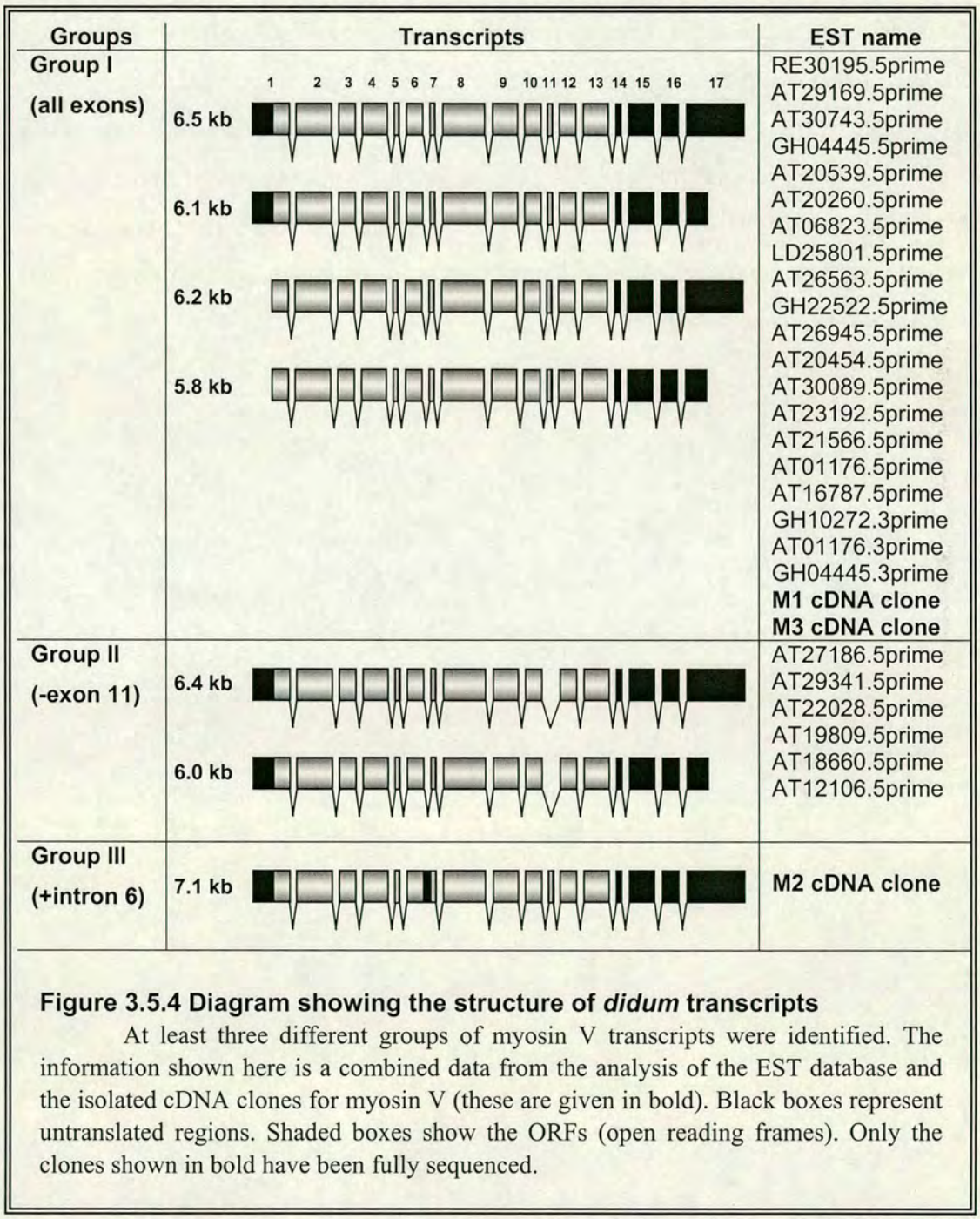


Figure 3.5.4 Diagram showing the structure of *didum* transcripts

At least three different groups of myosin V transcripts were identified. The information shown here is a combined data from the analysis of the EST database and the isolated cDNA clones for myosin V (these are given in bold). Black boxes represent untranslated regions. Shaded boxes show the ORFs (open reading frames). Only the clones shown in bold have been fully sequenced.

3.6 Discussion

We used a PCR based strategy to precisely map the *Drosophila* myosin V gene, *didum*, to region 43D1 on the polytene chromosome. A detailed analysis has shown that all known genes corresponding to position 43C-43D (position of saturated mutagenesis (Heitzler et al. 1993), are essential genes and lead to an embryonic lethal phenotype.

To further study *didum* we isolated cDNA clones and searched the EST database to elucidate the molecular structure of the gene. The analysis showed that the gene has a complex structure. It contains 17 exons that are spliced differentially to produce various transcripts expressed at different levels during the *Drosophila* lifecycle. The gene is transcribed from at least two putative promoters, thus producing transcripts with different 5'UTR regions. The 3'UTR contains two polyadenylation signals resulting in the production of two main groups of transcripts: long (6.5 kb) and short (6.0 kb) transcripts.

Chapter Four: Expression analysis of *Drosophila* Myosin V

4.1 Production of Myosin V antibodies

Recent studies have revealed that unconventional myosins perform a variety of essential cellular function such as organelle, RNA and protein transport, maintenance of the cell architecture, cell movements and signal transduction. Myosin Vs have been specifically implicated in vesicular and mRNA trafficking. *Drosophila* Myosin V contains a tail region very similar to the tails of other myosins from the same class. To gain insight into the function of *Drosophila* Myosin V we generated antibodies to a recombinant Myosin V from *Drosophila* and studied its subcellular localisation.

An antibody to *Drosophila* Myosin V have been previously generated in our lab (Bryce McIver, personal communications). Unfortunately this antibody was not able to recognise native Myosin V from *Drosophila* well. To check its specificity we performed a Dot-Blot Western analysis (Fig. 4.1A, see materials and methods, chapter 2.6.2). The Dot-Blot experiment showed that the generated antibodies to Myosin V bind with equal intensity to both the recombinant protein and to proteins from the nonexpressing Myosin V cell extract. This suggests that the antibodies can not be used to accurately detect Myosin V in *Drosophila*. We assumed that if the recombinant protein used for generating the antibodies has not been completely purified, this might have led to the production of small amount additional antibodies against the native *E. coli* proteins. When used for immunodetection these antibodies will generate unspecific staining and mask the Myosin V signal. To avoid the unspecific staining we blocked an aliquot of the antibodies with cell extract not expressing Myosin V (250µl serum with 50µl cell extract). This step was expected to deactivate the contaminating antibodies and to increase the signal to noise ratio in favour of Myosin V. The antibodies purified in this way were used for developmental Western analysis to detect native *Drosophila* protein (Fig. 4.1B).

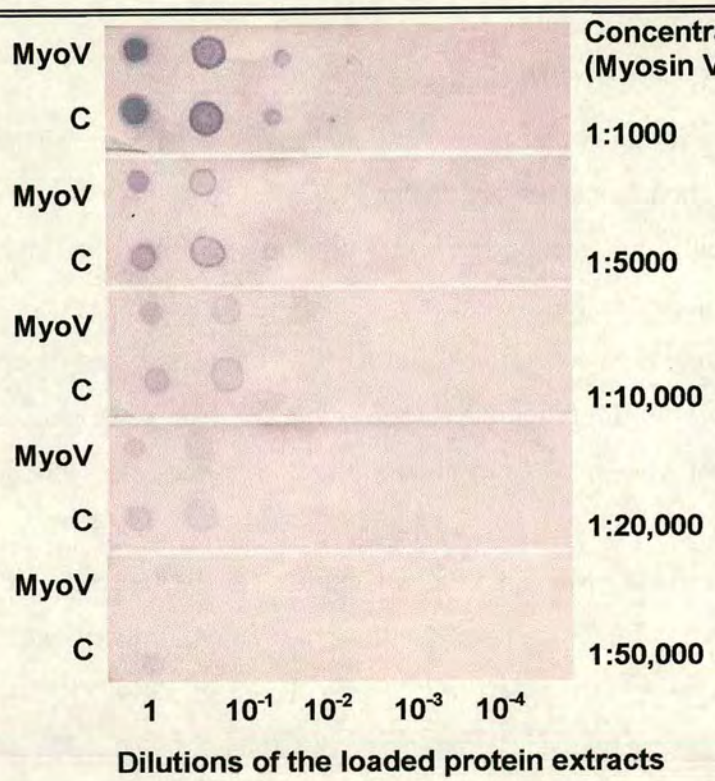


Figure 4.1A Western Dot-Blot analysis of the generated in rabbits Myosin V antibodies

Five nitrocellulose filters were spotted with ten fold dilutions (1μl each) of the recombinant protein (see materials and methods, chapter 2.6). The membranes were incubated with the Myosin V antibody with the shown concentrations. The concentration of the secondary antibody was constant (1:4000) for all five membranes. No difference was noted between the staining of the recombinant protein and the control *E. coli* proteins, suggesting that the antibodies bind unspecifically to cytoplasmic *E. coli* proteins. MyoV stays for Myosin V recombinant protein, C for control (extract from cells that do not contain expression vector and do not produce recombinant protein).

Samples from different developmental stages were collected and prepared for gel electrophoresis (see materials and methods, 2.6 chapter). The samples contain approximately the same amount of protein. The Western analysis showed that the preabsorbtion step has definitely improved the quality of the Myosin V antibodies. The antibodies were able to distinguish well the recombinant protein (80kDa) but were still binding to an unknown protein of approximately 45kDa (Fig. 4.1B). Unfortunately the antibodies were not able to detect the native Myosin V well. They produced a number of bands including a strong band of ~ 60 kDa from ovaries. It

should be noted that the antibodies did showed a diffuse band of approximately 200 kDa, which is the expected size for a single Myosin V heavy chain (the electrophoresis was carried out under denaturing conditions). This was especially strong in ovaries where the expression levels of Myosin V are known to be high.

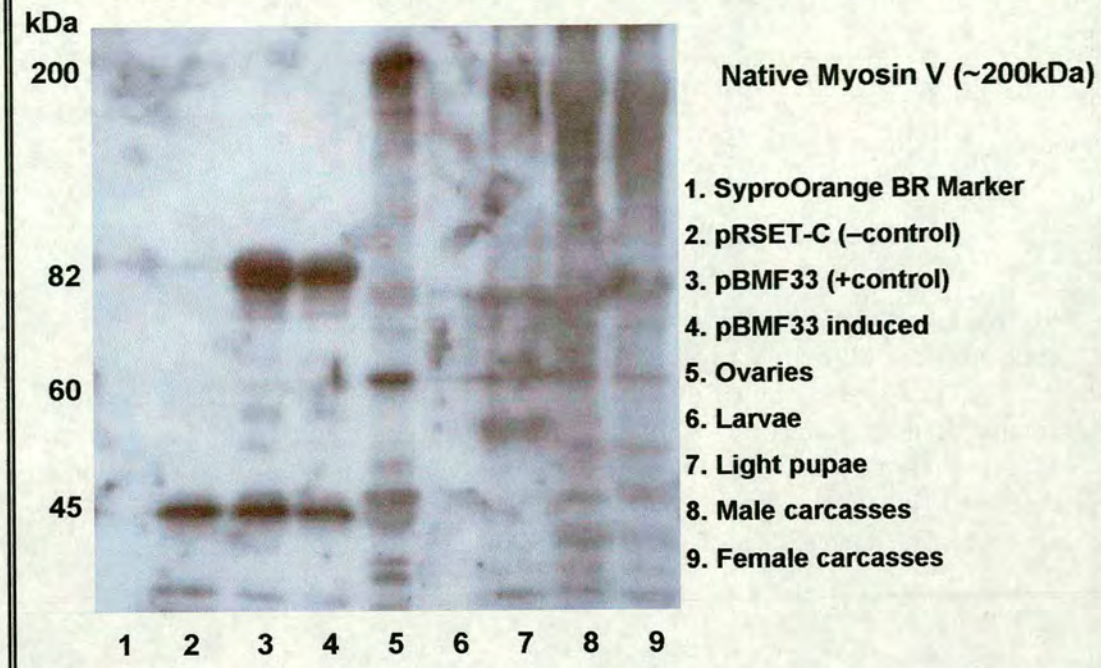


Figure 4.1B Developmental Western analysis of Myosin V expression

Immunochemical detection of native Myosin V in protein extracts from different developmental stages showed that the purified antibodies bind to native Myosin V (diffuse 200 kDa band) but also bind unspecifically to a number of other proteins. As controls cell extracts expressing and not expressing the recombinant protein were used. The molecular weight marker is SyproOrange, Broad range. It is not visible on the Western because the antibodies used does not recognise any of the proteins it containes.

4.1.1 Production and purification of recombinant Myosin V protein

4.1.1.1 Production of recombinant Myosin V protein

Myosin V antibody showed certain reactivity against *Drosophila* native protein. However further purification by affinity techniques was needed to increase its specificity. Since the antibody was generated in rabbits only 5ml of the immune serum were available after the preliminary experiments. Unfortunately this was not enough for further experiments. For this reason we decided to generate a second

batch of Myosin V antibody. This time the antibody was produced in sheep using the same recombinant protein as immunogen as the one used for the generation of the first antibody (Fig.4.1.1.1A).

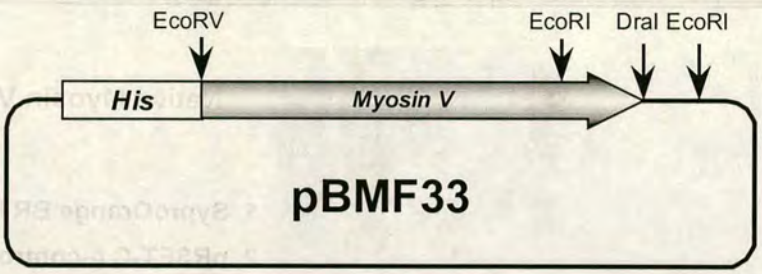


Figure 4.1.1.1A Expression vector used for the production of recombinant Myosin V protein

The vector pBMF33 was generated by Bryce McIver (personal communication). An *EcoRV-DraI* fragment (1986 bp) was excised from a *Drosophila* Myosin V cDNA clone. This fragment was cloned into pRSET-C expression vector at the *PvuII* site to produce an in-frame fusion with the N-terminal histidine tag sequence. The orientation of the insert was verified by *EcoRI* restriction.

Five stocks for pBMF33 (in BL21DE3-Lys) were stored from previous experiments. The cells were revived and grown to OD_{600}=1.3 and then induced with 1mM IPTG for 4 hours. Interestingly some of the stocks did not show any expression of recombinant protein (Fig. 4.1.1.1B).

To check the integrity of the expression vector we obtained plasmid DNA from these stocks. The DNAs were digested with *EcoRI* to check whether the studied clones are expressing the right protein. The digested samples produced a smear on agarose gels (data not shown), thus providing us with no information about the plasmid content. Some *E. coli* strains are known to express exonuclases. Since the cells expressing the recombinant protein showed the same smear on a gel when analysed, we assumed that this strain of BL1DE3-Lys also produces exonucleases. Thus if the obtained DNA is not perfectly purified, when digested with restriction endonucleases, the produced linear DNA molecules will become targets for the exonucleases. Unfortunately we did not find any relevant information about this in the literature. To check our hypothesis we transformed XL1Blue cells with the

expression plasmid DNA. Plasmid DNA obtained from p33XL (pBMF33 in XL1Blue cells) produced the correct restriction pattern. In addition we sequenced the plasmid to be sure that this is the correct expression vector. Interestingly p33XL cells showed no detectable levels of recombinant Myosin V expression.

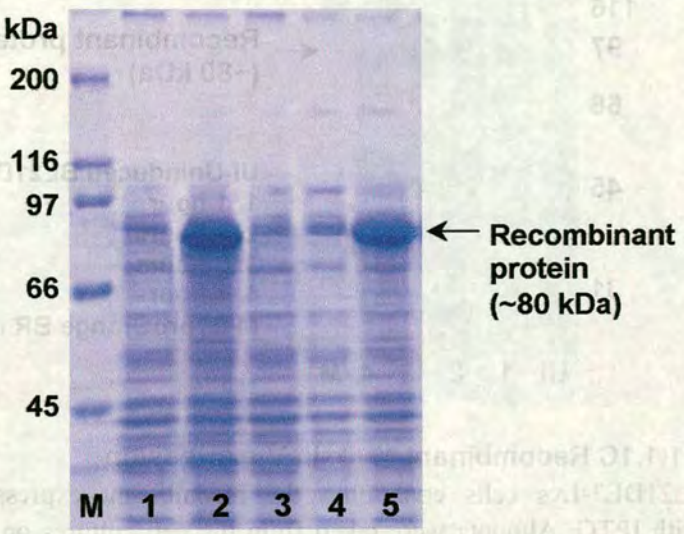


Figure 4.1.1.1B Recombinant protein expression

Five separate stocks of BL21DE3-Lys expressing the recombinant pBMF33 construct are shown. Cells from only stocks two and five expressed recombinant Myosin V. It was found later that the other three stocks also contain the expression construct but either the regulatory sequences of the vector or the cloned gene itself were damaged. The marker is SyproOrange, Broad Range.

We used plasmid DNA from the checked p33XL cell to transform back into BL21DE3-Lys strain. This strain is optimised for protein expression. It encodes T7 RNA polymerase under the control of *lacUV5* promoter. Induced with IPTG the promoter transcribes T7 RNA polymerase that is used to drive the expression of the recombinant protein (pRSET-C is *T7* promoter driven vector). In addition BL21DE3-Lys strain constitutively expresses low levels of T7 lysozyme which inhibits the promoter. This is especially useful to control the expression of toxic proteins or those damaging to the cells. When the cells are induced with IPTG they express high levels of T7 RNA polymerase which overcomes the lysozyme repression and the recombinant protein is produced.

One of new BL21DE3-Lys clones (called 2N33) expressing the recombinant Myosin V was selected. The cells were grown to OD_{600}=1.3 and induced with 1mM IPTG for one, two, three, and four hours (Fig. 4.1.1.1C).

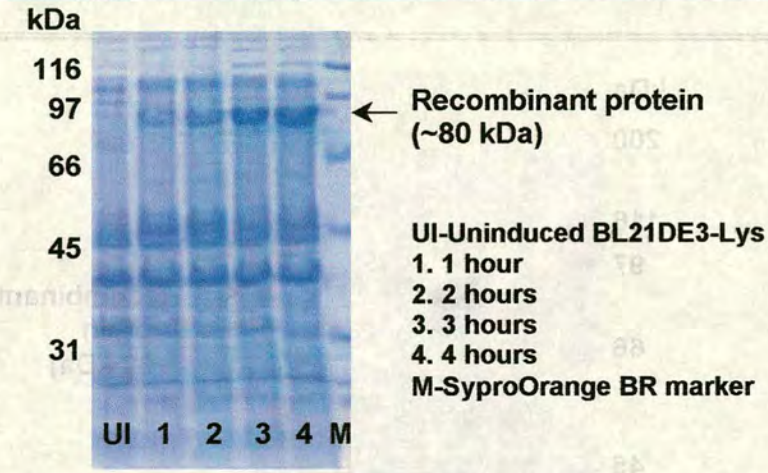


Figure 4.1.1.1C Recombinant Myosin V expression

BL21DE3-Lys cells containing the recombinant expression vector were induced with IPTG. Aliquots were taken from the cell cultures on the first, second, third and forth hour after the induction. The cell concentration was determined spectrophotometrically and the same amount of cells were taken for electrophoretical analysis. It was noted that after induction the cells almost stopped growing and switched to producing recombinant protein. Cells induced for four hours demonstrated the highest levels of expression and this time was set as the standard default in all subsequent experiments.

4.1.1.2 Affinity purification of recombinant Myosin V protein

To purify the protein we used Ni-NTA-Agarose. This resin allows the purification of His-tagged fusion proteins. In general the recombinant protein is adsorbed on a column prefilled with Ni-NTA-Agarose. The column is then washed and the protein eluted under high salt conditions.

Fractions II, III, and IV appeared to contain most of purified recombinant protein. To obtain enough protein we repeated the purification step two more times. Fractions II to IV from the three purifications were combined and dialysed twice to remove the elution buffer and to adjust the pH to physiological levels (pH~7.4). Approximately 14ml of dialysed protein were recovered. The protein solution was ten fold concentrated using Millipore concentration filters (UFV4BTK25).

The amount of recombinant protein recovered from the two affinity purifications was estimated to be 4.4 mg based on spectrophotometrical analysis. Approximately half of this protein was used to generate antibody in sheep.

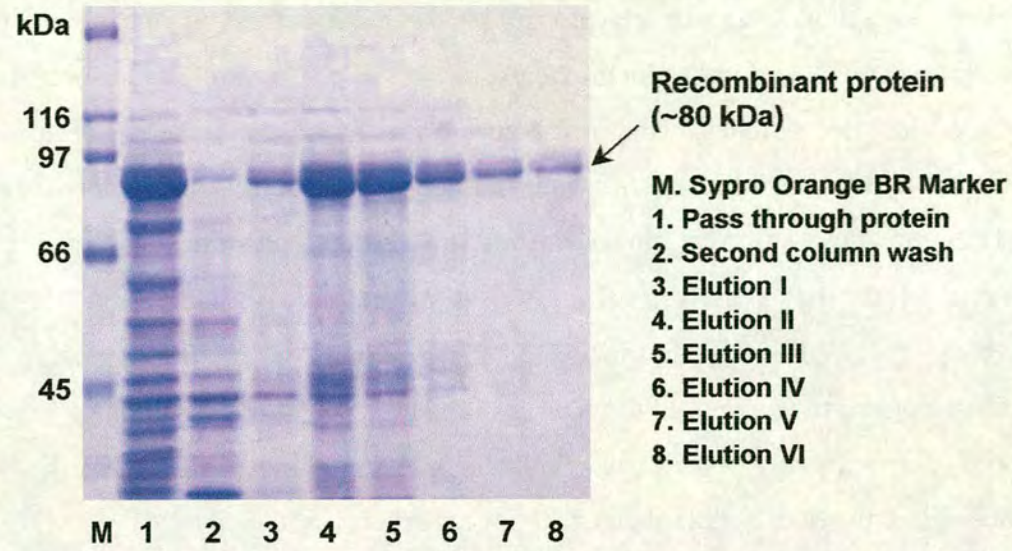


Figure 4.1.1.2 Affinity purification of recombinant Myosin V protein

A Ni-NTA-Agarose preloaded column was saturated with recombinant protein (1). Low concentration salt solution was used to wash the unbound protein (2). The protein was eluted with high concentration Imidazol buffer (see materials and methods). Ten fractions were collected, only six of which are shown on the figure. Most of the purified protein was in fractions II, III, and IV.

4.1.2 Sheep immunisation program

The antibody to *Drosophila* Myosin V was generated in sheep by Diagnostics Scotland (Scottish National Blood Transfusion Service). A pre-immune serum was isolated for control purposes. Then the animal (no: S256A) was immunised with 250μg of the recombinant protein. Three additional injections (250μg protein each) were carried out afterwards. The injections were spaced out approximately one month apart. Seven days after each injection serum was collected and processed (centrifuged to remove the cell debris) by staff members of Diagnostics Scotland.

4.2 *Affinity purification of Myosin V antibodies*

4.2.1 Preliminary Results with whole serum

Samples from the pre-immune serum and the first bleed were analysed by Western Dot-Blot. Cell extracts prepared from cells expressing the recombinant protein and cells that do not contain the expression vector (control) were spotted onto nitrocellulose membranes. The membranes were incubated either with the pre-immune serum or with serum from the first (immune) bleed. Both the pre-immune and the immune serum were blocked with a small amount of protein extract prepared from BL21DE3-Lys cells.10µl of serum were mixed with 20µl sonicated BL21DE3 cells containing pRSET-C and incubated for 40 minutes at room temperature. The antibodies (serums) were used for detection immediately after the blocking. This step would eliminate any unspecific binding of the antibody. The westerns were processed as previously described (materials and methods, chapter 2.6).

The results clearly show that the generated Myosin V antibody recognises the recombinant protein (Fig4.2.1C and D). When used at 1 to 20,000 time dilution of the sheep anti-Myosin V antibody produced a strong specific signal in comparison to the signal produced by the pre-immune serum (Fig.4.2.1D).

Serum obtained from the second and third bleed produced similar results (data not shown) demonstrating an adequate immune response in the organism in which they were generated. We decided to use the serum from the third bleed for our future experiments, because it showed a stronger signal when analysed on a Dot-Blot Western than serums from the other two bleeds.

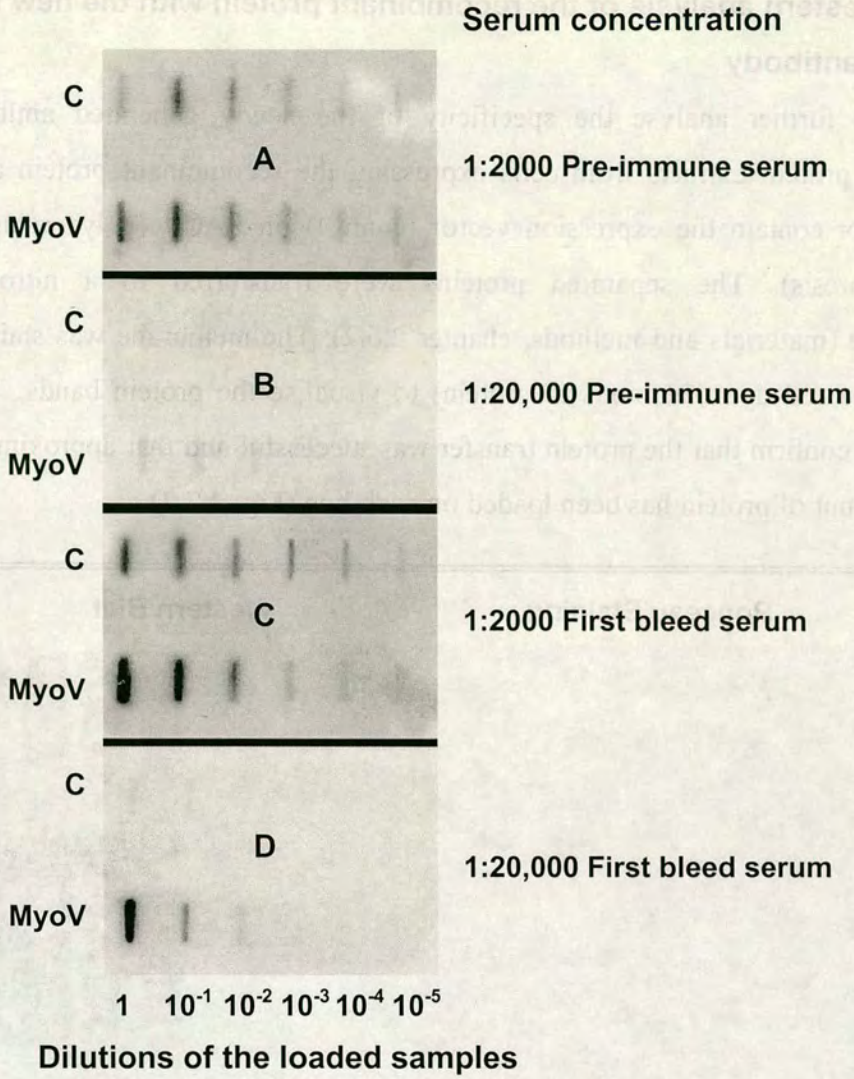


Figure 4.2.1 Western Dot-Blot analysis of the newly generated Myosin V antibodies

Four nitrocellulose filters were spotted with different dilution of the recombinant protein. Extract from cells that do not have the expression vector was used as a control. Two of the membranes were incubated in different dilutions of the pre-immune serum, and two in dilutions prepared from the immune first bleed. The concentration of the secondary (anti-sheep donkey IgG peroxidase-conjugated) antibody was constant (1:4000) for all four membranes. MyoV stays for Myosin V recombinant protein, C for control.

4.2.2 Western analysis of the recombinant protein with the new Myosin V antibody

To further analyse the specificity of the newly generated antibody we separated protein extracts from cells expressing the recombinant protein and cells that do not contain the expression vector (control) on PAGE (polyacrylamide gel electrophoresis). The separated proteins were transferred to a nitrocellulose membrane (materials and methods, chapter 2.6.2). The membrane was stained with Ponceau (a red dye that binds to protein) to visualise the protein bands. This was needed to confirm that the protein transfer was successful and that approximately the same amount of protein has been loaded on each line (Fig. 4.2.2).

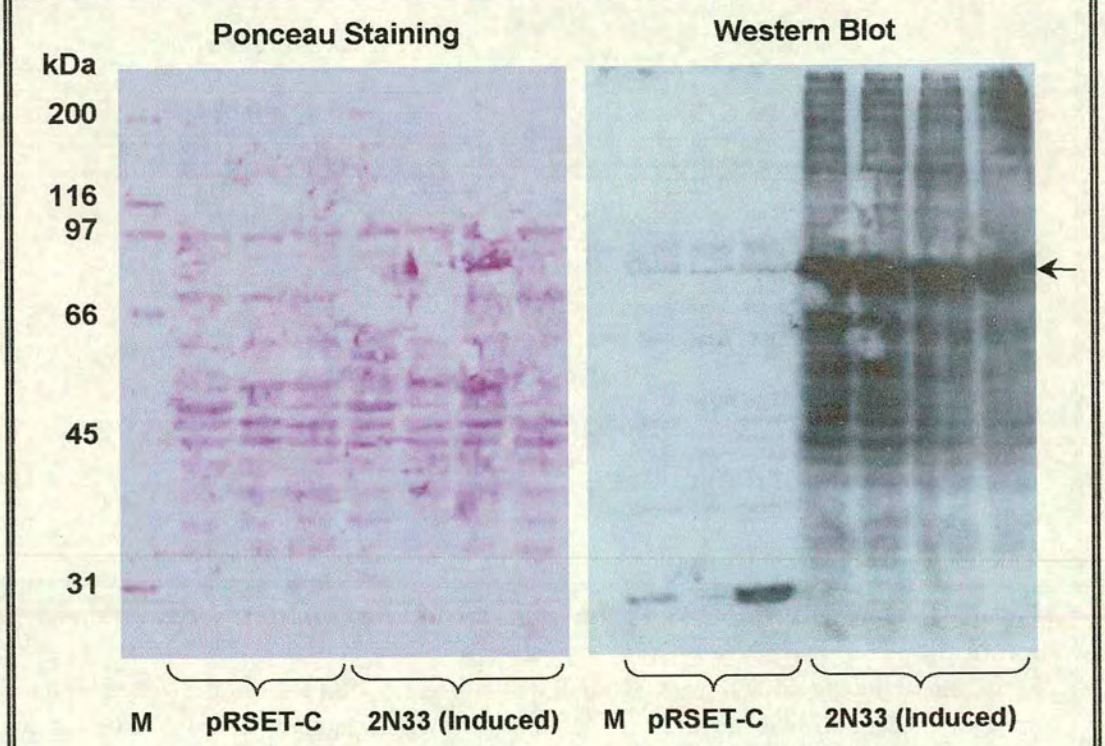


Figure 4.2.2 Western analysis of the recombinant protein with the new Myosin V antibody

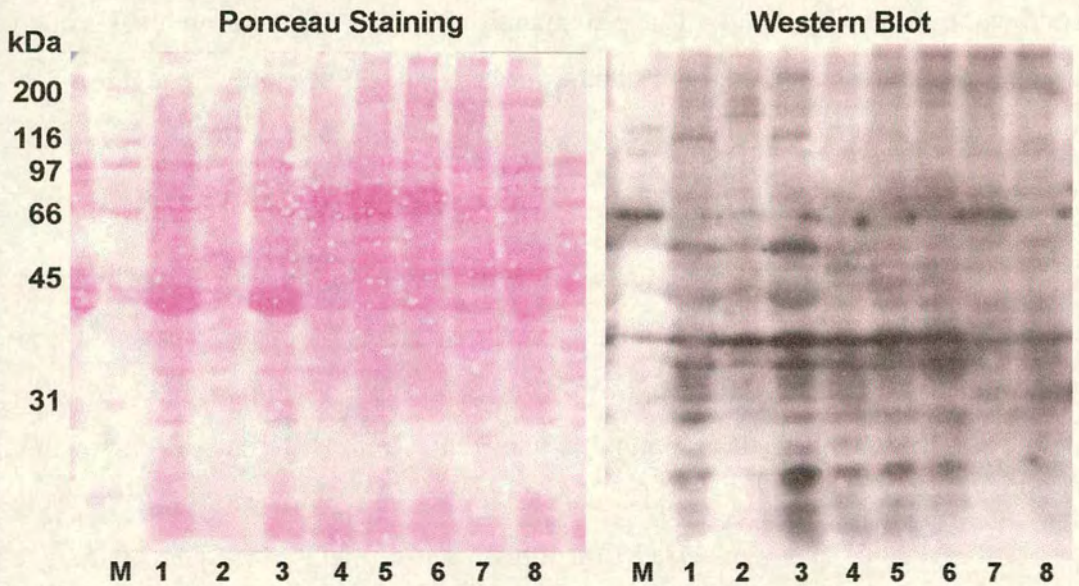
Immunochemical detection of the recombinant Myosin V in protein extracts from BL21DE3-Lys containing the expression 2N33 vector is shown. As a control a extract from BL21DE3-Lys with pRSET-C was used. The antibody detects a band of approximately 80 kDa (black arrow) only in the cells expressing the recombinant protein. The size corresponds to the recombinant protein. The marker is SyproOrange, Broad range. It is not visible on the Western because the antibodies did not recognise any of the proteins it contained.

The results showed that the new Myosin V antibody detects only the recombinant protein, a band of approximately 80 kDa. Almost no background staining was observed in the control lanes where there were only BL21DE3-Lys containing pRSET-C vector (but not the expressing recombinant vector).

4.2.3 Detection of wild-type Myosin V with the new Myosin V antibody

The results from the previous two experiments indicated that the new antibody might be capable of recognising the wild-type protein from *Drosophila* tissues. To check this we prepared protein extracts from different developmental stages (see materials and methods, chapter 2.6). The wild-type proteins were separated on PAGE and transferred to a nitrocellulose membrane. The membrane was blocked in Blocking buffer and the Myosin V antibody added at 1:20,000 dilution in 40ml TBS-T (one volume of antibody was pre-blocked for 30 minutes with three volumes of the cell extract prepared from BL21DE3-Lys). The membrane was then washed and blocked with 0.5% Donkey serum in TBS-T. This step was required to eliminate the unspecific binding of the secondary antibody to sites on the membrane. The secondary antibody (anti-sheep donkey IgG peroxidase-conjugated) was used in 1:4000 dilution.

The antibody did recognise the native Myosin V, a weak band of approximately 200 kDa (Fig. 4.2.3). However a strong background was present. This might have been because too much protein was loaded on the gel or because the antibody was not affinity purified and could bind non-specifically. To check the first possibility we repeated the Western analysis with lower amount of protein. Almost no difference was observed when compared to the previous Western (results not shown).



M. SyproOrange BR Marker
1. Ovaries
2. Testes
3. Embryos
4. Larvae (IIIrd instar)
5. Light pupae
6. Dark pupae
7. Male carcasses
8. Female carcasses

Figure 4.2.3 Western blot with *Drosophila* samples

Protein extracts from different developmental stages are analysed. The Ponceau staining indicates that protein has been loaded equally in each lane. The new Myosin V antibody detects a high molecular weight protein around 200 kDa corresponding well with the expected size for native *Drosophila* Myosin V (marked with an arrow). The antibody binds also to proteins with lower molecular weight. These might be either different Myosin V isoforms (produced by alternative transcripts), or unspecific binding to other proteins containing domains similar to the myosin tail.

4.2.4 Caprylic acid purification of the crude serum

Caprylic Acid, at the right concentration and at a particular pH, will precipitate most serum proteins except IgG. This step alone often yields more than 85% purity. Crude serum (third bleed) was processed as described in chapter 2.6.7.1. The purified antibody solution was concentrated ten fold using UFV4BTK25 filters from Millipore (Fig. 4.2.4).

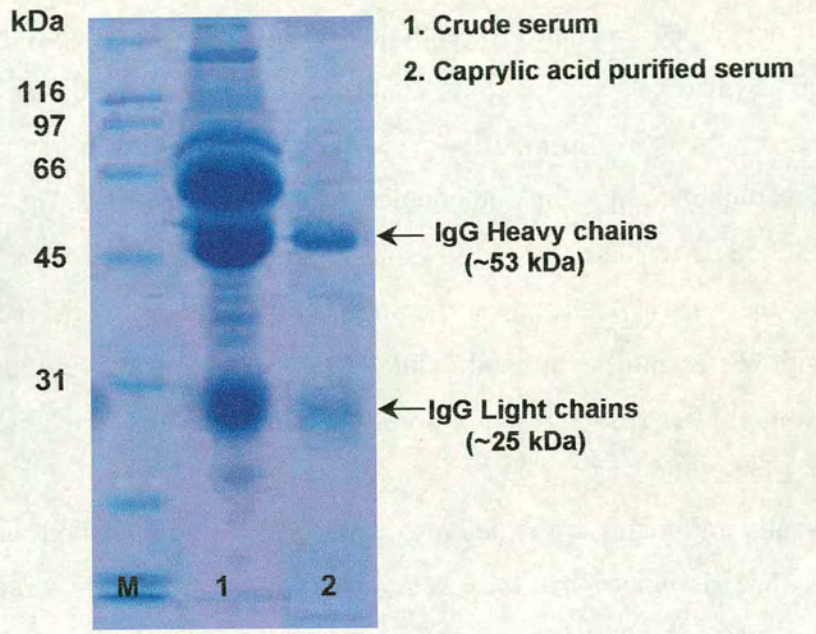


Figure 4.2.4 Caprylic acid purification on Myosin V antibody

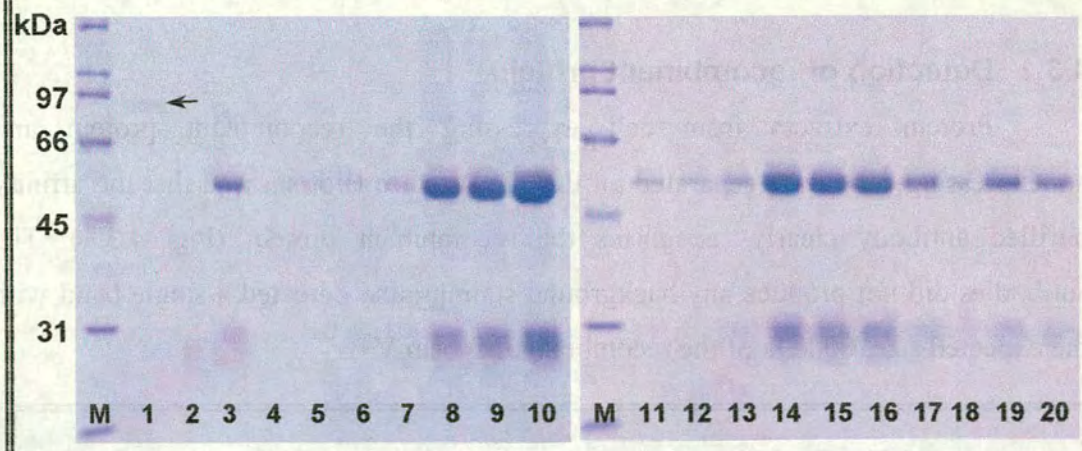
The polyclonal antiserum was subjected to caprylic acid precipitation. This removed the bulk of extraneous serum proteins such as albumins. The figure shows the two chains of IgGs dissociated from each other (denaturing PAGE).

4.2.5 Affinity purification of Myosin V antibody

To further purify the immunoglobulin fraction it was subjected to affinity chromatography using Affigel beads (Bio-Rad). Approximately 2 mg of recombinant protein was coupled to Affigel beads 10 and processed in accordance with the manufacturer's instruction. Unfortunately no antibody was recovered using this method. PAGE analyses showed that no recombinant protein binds to the Affigel beads 10.

We developed a method for antibody affinity chromatography based on Ni-NTA Agarose (materials and methods, chapter 2.6.7.2). In previous experiments Ni-NTA Agarose was shown to bind the recombinant protein. We used approximately 0.5 mg of recombinant Myosin V to couple with Ni-NTA Agarose. The caprylic acid purified antibody was then allowed to bind to the recombinant protein. The column containing the Ni-NTA Agarose-Recombinant protein-Antibody complex was washed with buffer and the antibody eluted. We carried out two elutions, one with acidic Glycine buffer (pH-2.5) and a second one with a basic Triethylamine buffer (pH-11.5) (Fig4.2.5).

Usually antibodies are eluted at specific pH. In our case both buffers eluted antibodies. We assumed that the each buffer eluted a specific fraction of the polyclonal antibody. The two fractions were combined for the subsequent experiments (Fig.4.2.5). The combined fractions were dialysed against PBS and concentrated 10 fold on UFV4BTK25, Ultrafree-4 centrifugal filters from Millipore.



M. (SyproOrange) BR Marker
1. Diluted Protein (arrow) before loading on the column
2. Diluted Protein after loading on the column
3. Diluted Antibody before loading on the column
4. Diluted Antibody after loading on the column
5. Salt wash I
6. Salt wash II
7. Salt wash III
8. Glycine elution I
9. Glycine elution II
10. Glycine elution III

M. (SyproOrange) BR Marker
11. Tris HCl, pH-8.8 wash I
12. Tris HCl, pH-8.8 wash II
13. Tris HCl, pH-8.8 wash III
14. Triethylamine, pH-11.5 elution I
15.Triethylamine, pH-11.5 elution II
16. Triethylamine, pH-11.5 elution III
17. Purified antibody, 50x dilution
18. Purified antibody, 300x dilution
19. Concentrated antibody, 500x dilution
20. Concentrated antibody, 1.10^{-3} dilution

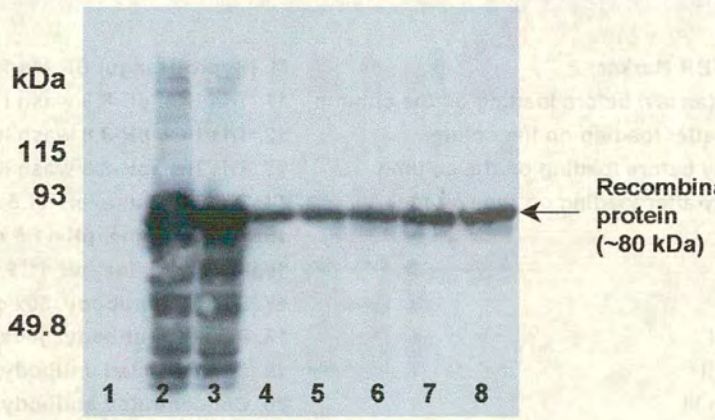
Figure 4.2.5 Affinity purification of Myosin V antibody

To purify the antibody we used a method based on Ni-NTA Agarose. The separate steps of the purification were monitored by sampling. The antibody-containing fractions 8,9,10,14,15 and 16 were combined and used for further experiments.

4.3 *Western analysis of Myosin V antibody*

4.3.1 Detection of recombinant protein

Protein extracts from cell expressing the recombinant protein and nonexpressing cells were separated on PAGE. Western Blot showed that the affinity purified antibody clearly recognises the recombinant protein (Fig. 4.3.1). The antibodies did not produce any background staining and detected a single band with the expected size, 80 kDa of the recombinant Myosin V.

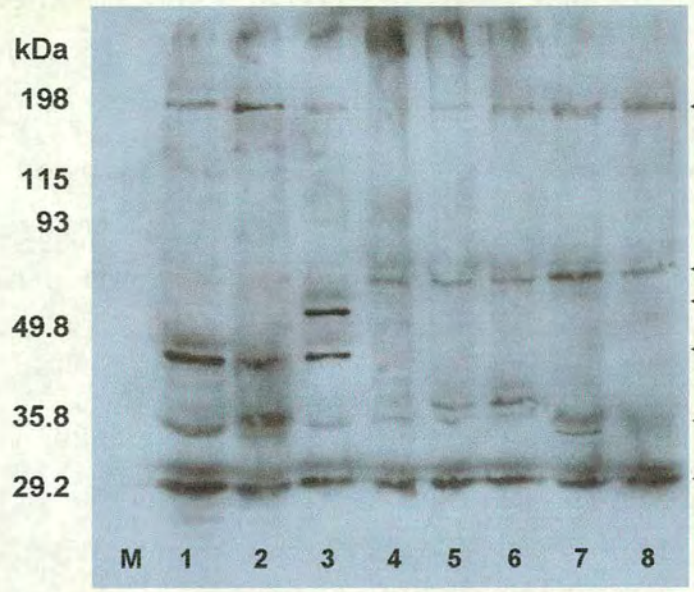


1. pRSET-C
2. p33XL-1 (overloaded)
3. p33XL-2 (overloaded)
4. 2N33, non-induced
5. 2N33, induced (1hour)
6. 2N33, induced (2 hours)
7. 2N33, induced (3 hours)
8. 2N33, induced (4 hours)

Figure 4.3.1 Western blot with recombinant protein
The antibody used was affinity purified from a Ni-NTA Agarose column. The Marker is a SyproOrange, Broad range. It is not shown on the picture because it is not visible on Westerns-the antibodies do not recognise any of the proteins it contained.

4.3.2 Detection of wild-type Myosin V

Extracts from different developmental stages were analysed on a Western Blot (Fig. 4.3.2). This time the antibody produced much less background and was clearly able to detect the 200 kDa of *Drosophila* wild-type Myosin V.



M. (SyproOrange) BR Marker

1. Ovaries
2. Testis
3. Embryos
4. Larvae (IIIrd instar)
5. Light pupae
6. Dark pupae
7. Male carcasses
8. Female carcasses

Figure 4.3.2 Western blot with *Drosophila* extracts

Protein extracts from different developmental stages are analysed. The new Myosin V antibody detects a high molecular weight protein, around 200 kDa, corresponding well with the expected size for native *Drosophila* Myosin V (marked with an arrow). The antibody also binds to proteins with lower molecular weight. These might be either different Myosin V isoforms (produced by alternative transcripts), or unspecific binding to other proteins containing domains similar to the myosin tail used to generate the antibody.

Currently we do not have explanation why the antibody binds to other proteins and produces multiple bands on the Western Blot. It seem that this is unspecific binding because when Western Blot is done with protein extracted from embryos, a single stage where the conditions (blocking time, amounts of antibody used, temperature) could be optimised easily, the result is clear Blot showing only two bands, a strong 200kDa and faint band at 40kDa (see results in chapter 5).

4.4 *Immunolocalisation of Myosin V in Drosophila ovaries*

To study the subcellular distribution of Myosin V during *Drosophila* oogenesis we used both whole serum and affinity purified antibodies. The method is described in materials and methods, chapter 2.6.3. We used fluorescent secondary antibodies which were later visualised and photographed on fluorescent microscope. To show the precise localisation of Myosin V we performed triple staining experiments.

4.4.1 Myosin V/Actin and non-muscle Myosin II/Actin staining

Myosins bind and move along actin filaments. Studies have shown that non-muscle Myosin II is expressed during oogenesis (Edwards and Kiehart 1996). We know from our Western experiments that Myosin V is also produced during *Drosophila* oogenesis. Immunostaining of *Drosophila* ovaries showed that non-muscle Myosin II is expressed abundantly through out the oogenesis. During stages 1-10 the protein is detected in the somatically derived follicle cells, as well as in the germ-line nurse cells (Fig. 4.4.1A, C). The protein is also seen in the cells of the epithelial sheet overlaying the developing egg chambers. The levels of Myosin II decrease with the advancing of the oogenesis. In later stages, 11-13 there is less protein in the degenerating nurse cells and the protein in the follicle cells from being diffuse becomes localised to the cell membrane (but not the nuclear membrane) (Fig. 4.4.1E, G).

Actin is a major cytoskeletal protein and is expressed at all stages of oogenesis. It localises mainly to and closely to the cell membrane and the diffuse immunostaining indicates its presence throughout the cell cytoplasm. Generally Myosin II co-localises with Actin. However Actin is a main constituent of the ring canals (Fig. 4.4.1C-yellow triangles) while Myosin II is not. In addition Actin is more strongly expressed in the periphery of the oocyte at the base of the follicle cells (Fig. 4.4.1C-yellow arrows).

Myosin V is mainly localised in the nurse cells at stages 1-13 from where it is likely to be transferred to the growing oocyte - a weak staining is observed in the oocyte at stage 1-10. Expression was also observed in the follicle cells which was not detected by *in situ* hybridisation to mRNA, presumably due to low transcripts levels.

At stage 10-12 Myosin V was observed in a region surrounding the border cells, but not in the border cells themselves (Fig. 4.1.1D). At this stage the columnar follicle cells start to migrate centripetally to cover the anterior end of the oocyte (Fig. 4.1.1D, F). Since the oocyte remains connected to the nurse cells by the ring canal, the latter becomes localised centrally being surrounded by the migrating follicle cells. Our observations show that Myosin V staining is localised to that part of the ring canals. In addition Myosin V staining is detected in the anterior part of the oocyte close to ring canal opening. In later stages,10-13, a dynamic pattern was observed in the columnar follicle cells but not in the stretched follicle cells. Myosin V staining is confined to the cell cytoplasm (Fig. 4.4.2D) and almost no staining is present in proximity to the cell membrane, compared to the expression of Myosin II at these stages (Fig4.1.1G, H).

Figure 4.4.1 Myosin V/Myosin II expression during *Drosophila* oogenesis

Antibody staining to Myosin II compared to Myosin V is shown. Both Myosin V and Myosin II have been detected using FITC labelled antibodies (Green); Actin is labelled with TRITC labelled phaloidine (Red); DNA content is labelled with Hoechst (Blue).

A: Immunostaining with specific antibodies to Myosin II revealed its presence throughout the oogenesis. The protein is expressed abundantly in both nurse and follicle cells during the early stages of oogenesis. Its co-localisation with Actin is shown.

B: Myosin V is expressed in a pattern similar to that produced by Myosin II, however the expression levels are lower.

C: Localisation of Myosin II at stage 10 of oogenesis. Myosin II is strongly expressed in the nurse cells and all types of follicle cells. Actin ring canal localisation is indicated with yellow triangles. High levels of Actin expression are also detected in the oocyte just under the follicle cells (yellow arrows).

D: Dynamic staining of Myosin V at stage 10. The protein is localised in the ring canal connecting the nurse cells with the oocyte (white arrows). Diffuse staining indicates the presence of Myosin V in the anterior part of the oocyte.

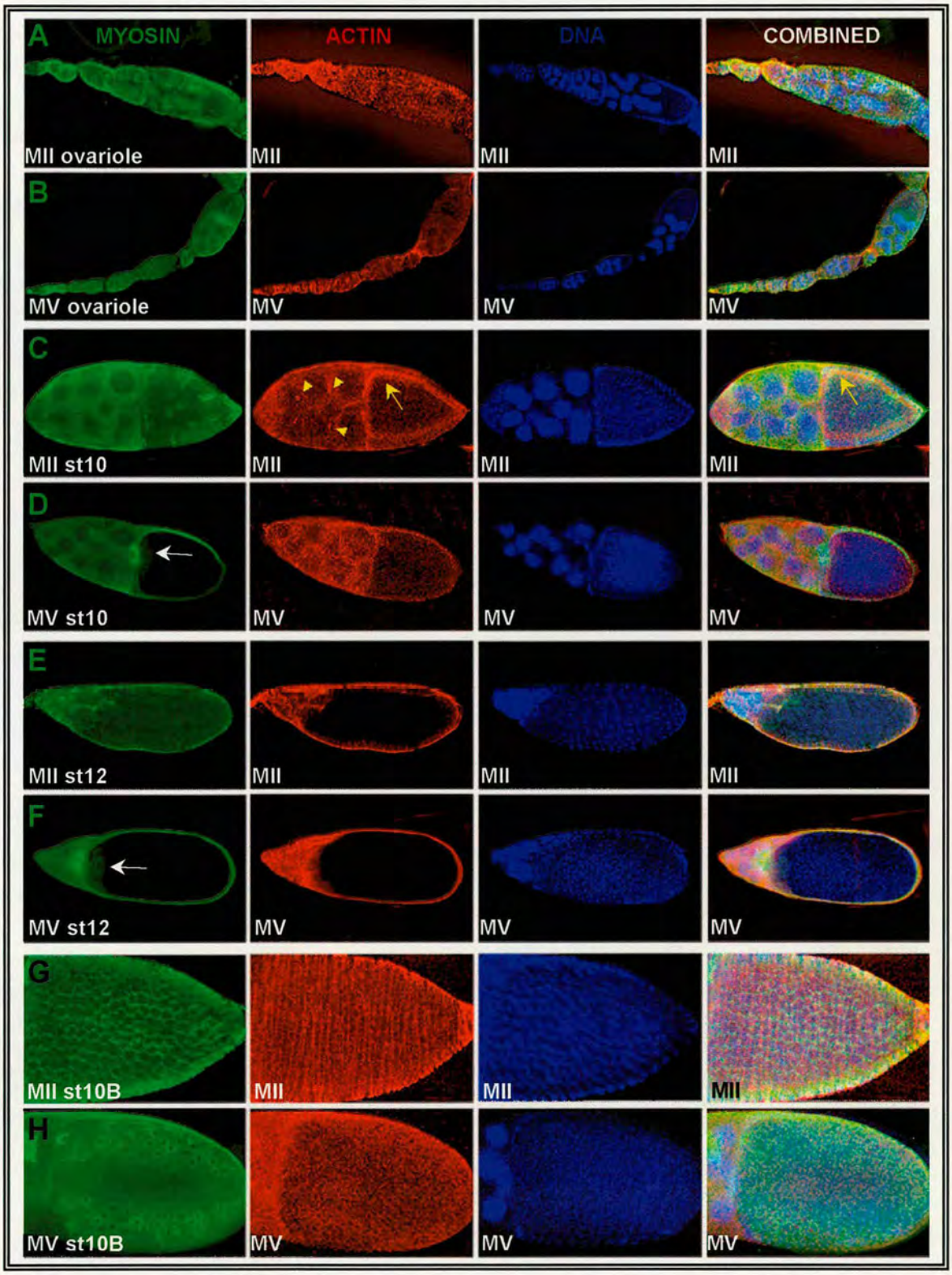
E: Low levels of Myosin II expression are detected in the follicle cells and degenerating nurse cells at stage 12.

F: The localisation of Myosin V protein in the ring canal and the anterior part of the oocyte is still detectable at stages 11-13 of oogenesis.

G: High magnification of egg chamber stage 10. Myosin II is localised to the follicle cell membrane.

H: Dynamic expression of Myosin V in the follicle cells cytoplasm, stage 10. ⟶

Figure 4.4.1 Myosin V/Myosin II expression during *Drosophila* oogenesis



4.4.2 Myosin V/Calmodulin staining

Myosin V has a long neck domain with six IQ motifs. It has been shown that at least four of the IQ motifs bind to calmodulin. The other two bind additional 25kDa and a 17kDa light chains (Chenney et al. 1993). A dimer molecule should associate with eight calmodulin subunits. This suggests that Myosin V and calmodulin should co-localise. In this way we can check the specificity of our Myosin V antibody. To study the distribution pattern of Myosin V compared to calmodulin we performed a double staining using specific antibodies.

The results from the immunostaining confirmed our hypothesis. During stages 1-10 the expression pattern of calmodulin closely matched the Myosin V expression pattern (Fig. 4.4.2A, B). The calmodulin signal like the Myosin V signal was confined to the same region at stage 10. The calmodulin was detected in the ring canal that links the nurse cell with the oocyte and in the anterior part of the oocyte (Fig. 4.4.2B, C, yellow arrows). At stage 10-13, the cytoplasmic localisation of calmodulin in the follicle cells did not match the dynamic Myosin V pattern. This is clearly seen when the two expression patterns were overlaid.

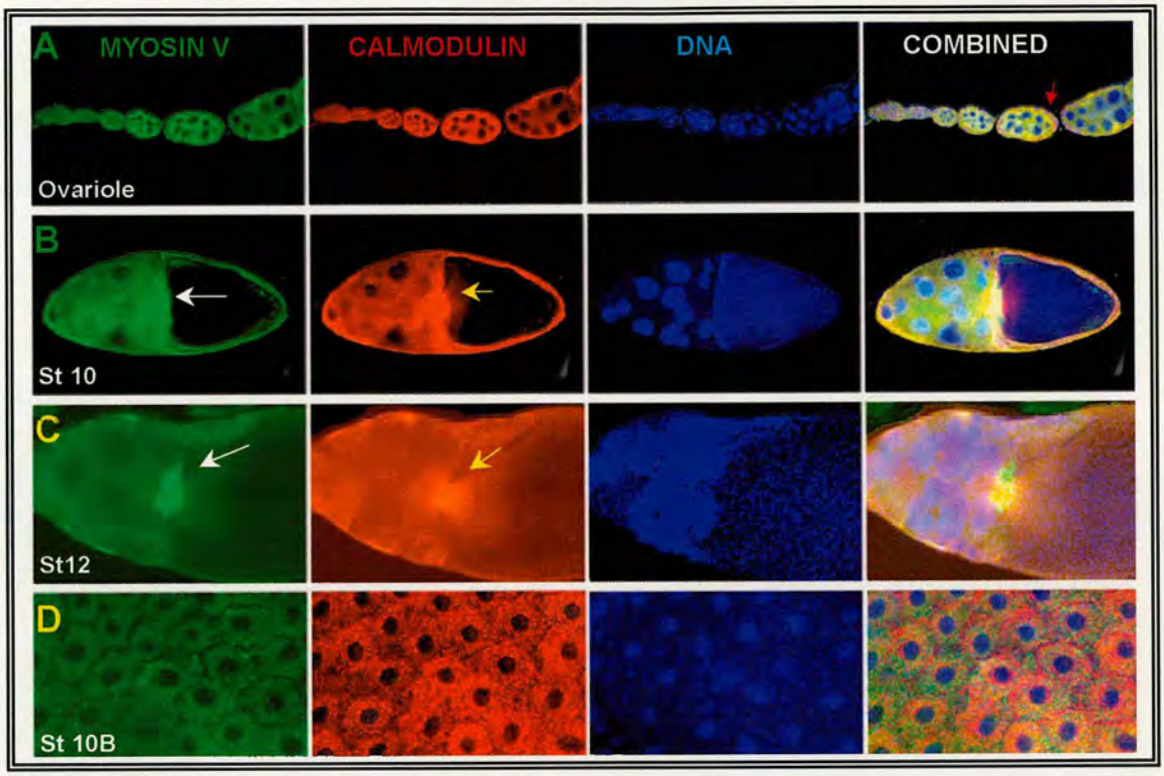


Figure 4.4.2 Immunolocalisation of Myosin V/Calmodulin during *Drosophila* oogenesis

Antibody staining to Myosin V versus calmodulin is shown. Myosin V has been detected using FITC labelled antibodies (Green); calmodulin is labelled with TRITC labelled antibodies (Red); DNA content is labelled with Hoechst (Blue).

A: The early expression patterns of calmodulin and Myosin V match closely. In addition, the calmodulin is localised to the perinuclear region. This produces characteristic red dots when the two images are overlaid (red arrow).

B: Expression pattern at stage 10 of oogenesis. Myosin V (white arrows) and calmodulin (yellow arrows) localise to the ring canal connecting the nurse cells with the developing oocyte; both proteins are delivered to the oocyte (anterior staining).

C: The characteristic ring canal pattern is maintained until stage 13 of oogenesis

D: High magnification of Myosin V and calmodulin expression patterns in the follicle cells, stages 10-13. Both proteins show a cytoplasmic distribution. Interestingly when overlaid they do not co-localise. If they did the green and red layers should have produce a yellow image as was the case in B and C.

4.5 Discussion

Antibodies to *Drosophila* Myosin V have been generated in sheep. The antibodies have been purified by affinity chromatography to increase their specificity. Western Blot experiments have shown that both crude serum and purified antibodies recognise the recombinant protein which was used as immunogen. The affinity purified antibodies were used to detect wild-type *Drosophila* protein and to build a developmental profile of Myosin V expression. Immunolocalisation experiments demonstrated that Myosin V is localised in the nurse cells at stages 1-13, from where it is transferred to the oocyte. A characteristic staining pattern was observed at stage 10 when the protein is seen in the region where the ring canal connects to the developing oocyte. At later stages, 10-13, a dynamic pattern was observed in the columnar follicle cells, where the protein is localised to the cell cytoplasm. The correlation between cell distribution and function will be discussed in details in chapter 5.

Chapter Five: Genetical analysis of Myosin V function

5.1 Search for mutations in didum

5.1.1 Search for P-element lines

P-element lines with a P-element inserted in the coding sequence of a gene of interest can be mutant for this gene. If no P-insertions are known for a given gene of interest a P-hop mutagenesis can be undertaken to create mutants (see the introduction chapter, 1.2.3).

We initially chose to study Line H14 because this was the original enhancer trap line which led to the discovery of *myosin V* because of a co-ligation. The line gave an interesting and dynamic expression pattern in the follicle cells. *In situ* hybridisation to the polytene chromosomes showed two P-element insertions (MacIver et al. 1998; A. McCormack Thesis, 1998) in region 41-43. It was possible that this was close enough to undertake a P-hop mutagenesis to generate a *myosin V* mutant. For this purpose it was decided to separate the two P-elements by recombination. Thus we generated A27 and A35 lines (McCormack Thesis, 1998). The two lines contained only one P-element but unfortunately it was the same one. To see if this P-insert is preferentially inserted we analysed the expression patterns of the reporter gene of the original (H14) and derived lines (A27, A35) (Fig. 5.1.1A). The reporter gene was expressed in subsets of follicle cells throughout oogenesis, except at stages 2-5. A27 and A35 showed an identical expression pattern that was similar to the one observed in H14. The expression pattern was highly interesting. Initially the reporter gene is expressed in the germarium. In later stages 6-12 a dynamic pattern is observed in the columnar follicle cells and the border cells, but not the stretched follicle cells. A27 and A35 expression pattern differ from H14 only in that that not all columnar follicle cells are stained at late stages, 10-12. Our analysis confirmed that the original H14 line has two P-elements both of which map to novel genes. The Inverse PCR experiments showed that they were inserted in region 41E and 42E which is relatively far from *didum* (43Da) to be used for P-hop mutagenesis. Lines H14, A27 and A35 were used for further analyses. Studying a gene with such a dynamic expression pattern can contribute to our knowledge of cell

interaction occurring between the germline cells and the somatically derived follicle cells.

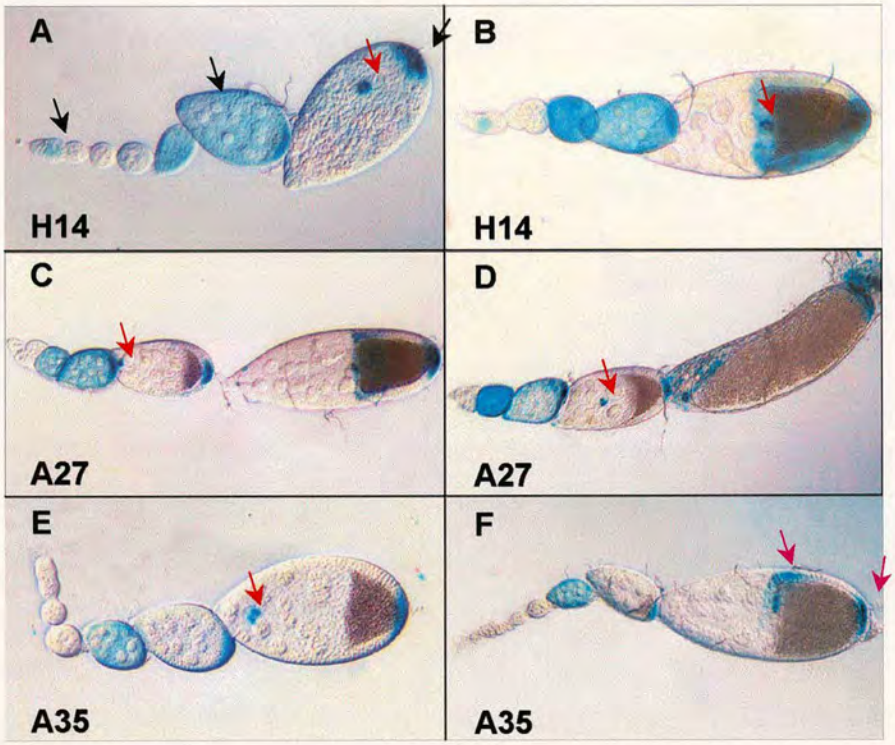


Figure 5.1.1A Reporter gene expression in lines H14, A27 and A35

Line H14 contains two P-elements. Lines A27 and A35 are identical and contain only one of the H14 P-elements.

A, B: Reporter gene expression in the parental H14 line. The gene is expressed in the germarium (black arrow). During stage 2-5 no staining is observed. At stages 6-7 the reporter gene is expressed in the follicle cell layer that covers the egg chamber. At stage 8-9 the staining is confined to the posterior follicle cells (black arrow). Staining is detected in all columnar follicle cells at stages 10-13. Strong expression is observed in the migrating border cells (red arrow)

C, D, E, F: A27 and A35 show identical staining patterns. This pattern is similar to the one observed in H14 where the pattern is a result of combined expression of two reporter genes. A27 and A35 expression pattern differ from H14 only in that that not all columnar follicle cells are stained at late stages, 10-12. The staining is confined to two patches of follicle cells: the posterior-most follicle cells and group of anterior follicle cells (purple arrows). A strong expression in the migrating border cells is also observed (red arrow).

A search of the available fly stocks, identified 14 lines with insertions in the 43B-E region (Table 5.1.1A). These were analysed by Inverse PCR to recover (rescue) the genomic sequence flanking the P-element (Fig. 5.1.1B, C, D). The recovered genomic sequences were sequenced to determine the position of the P-element insertion (materials and methods, chapter 2.5.11).

Table 5.1.1A P-element lines used in the screen for transposone insertion in *didum*

Stock no	Genotype	Chr*	Reference
H14	*P[lacZ, ry$^+$] ; P[lacZ, ry$^+$]*	2	MacIver et al. (1998)
A27	*P[lacZ, ry$^+$]*	2	MacIver et al. (1998)
A35	*P[lacZ, ry$^+$]*	2	MacIver et al. (1998)
BL-11412 or P1412	*P{ry+t7.2=PZ}Dscam05518 cn[1]/CyO; ry[506]*	2;3	Spradling et al. (1999)
k16128	*y1 w67c23; P{w+mC=lacW}cosk16128/CyO*	1;2	Torok et al. (1993)
BL-11156 or k16101 or k16122	*y1 w67c23; P{w+mC=lacW}cosk16101/CyO*	1;2	Torok et al. (1993)
BL-11382 or 04614	*P{ry+t7.2=PZ}l(2)43Bb04614a cn[1]/CyO, nec[JR]; ry[506] P{PZ}04614b*	2;3	Spradling et al. (1999)
k13522	*Df(1)w[67]c[23], y[1]; P{lacW}l(2)43Bd [k13522] / In(2LR)O, Cy dp[lvl] pr cn[2]*	2	John Root. (1995)
654 or ms (3)00919	*P{ry+t7.2=PZ}ms(2)43C00919/CyO; ry[506]*	2;3	Castrillon et al. (1993)
BL-10818 or P818 or 08815	*y1 w[67]c[23]; P{w+mC=lacW}dpldk08815/CyO*	1;2	Spradling et al. (1999)
11342 or P1342	*P{ry+t7.2=PZ}Aldh-III03610 cn[1]/CyO; ry[506]*	2;3	Spradling et al. (1999)
05467	*P{ ry+t7.2=PZ}l(2)03610[05467]*	2	Spradling et al. (1999)
4403 or k08255	*y1 w[67]c[23]; P{w+mC=lacW}l(2)k08255k08255/CyO*	1;2	Kania et al. (1995)
10467 or P467	*y1 w[67]c[23]; P{w+mC=lacW}l(2)k00107k00107/CyO*	1;2	Spradling et al. (1999)

* Chr for Chromosome. Specifies the chromosome position of the P-element insertion.

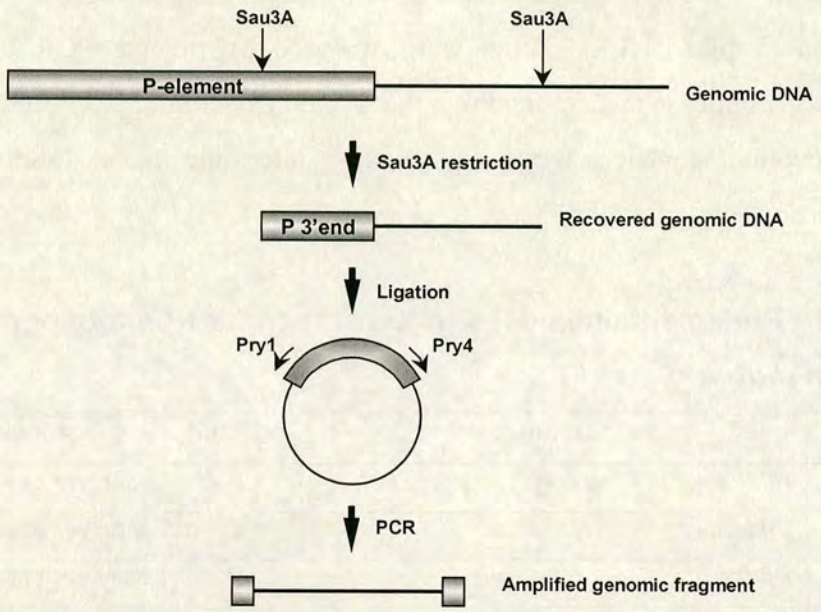


Figure 5.1.1B Principle of Inverse PCR

Described is 3' Inverse PCR. The techique of 5' Inverse PCR is similar. Genomic DNA is isolated from the target fly strain and digested with *Sau3A* restriction endonuclease. This enzyme has a site close to the 3' end of the P-element. In general *Sau3A* cuts approximately every 300-400 bp, thus there is a very high probability that there will be a *Sau3A* site close to the P-element in the flanking genomic DNA. The produced linear fragments are ligated with *T4DNA* ligase. To recover the genomic DNA adjacent to the P-element a PCR is performed. Subsequently the rescued sequence is

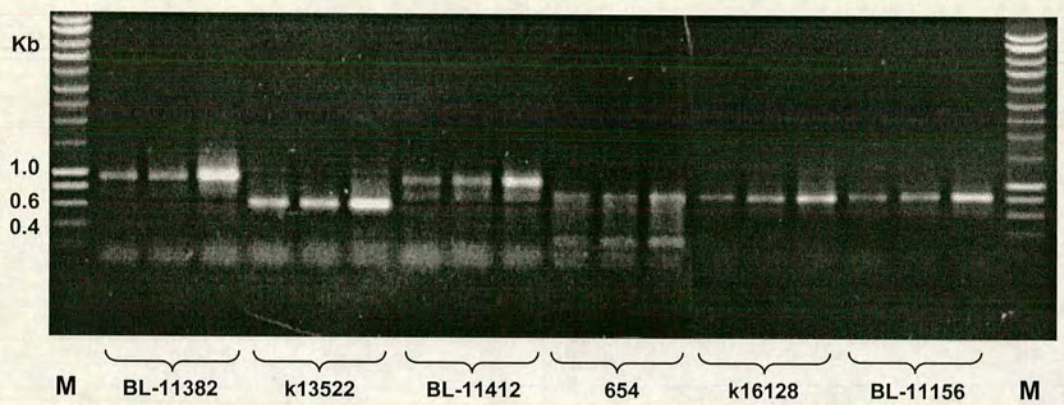


Figure 5.1.1C Inverse PCR

The PCR was performed at three different temperatures to chose the best conditions for amplification. The bands were excised from gel, purified and sequenced. The two faint bands produced for line 654 were used as templates in a second round of PCR to amplify sufficient DNA for sequencing. The marker used is HyperLadder (Bioline).

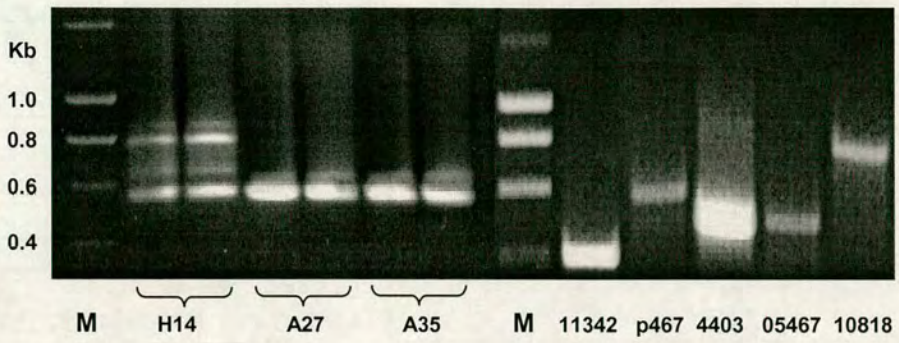


Figure 5.1.1D Inverse PCR

The PCR was performed at two different temperatures to chose the best conditions for amplification of H14, A27 and A35. The bands were excised from gel, purified and sequenced. Line H14 produces two bands because it contains two p-elements. Lines A27 and A35 derive from H14 and contain only one of the two P-elements. The marker used is HyperLadder (Bioline).

To determined the position and orientation of the inserted P-elements we used the sequenced DNA fragments to search the NCBI BLAST database (Table 5.1.1B, Fig. 5.1.1E). It was found that some of the lines contain two P-elements. In 654 the second P-element was located on chromosome 3, while in H14 it was inserted near the first P-element. k13522 had a single P-element inserted on the third chromosome.

The results from the Inverse PCR showed that none of the studied lines have a P-element in *didum* and are not mutants of this gene. In addition the precise mapping of the P-insertions revealed that all of them are too far away from *didum* to be useful for P-hop mutagenesis (see the introduction, chapter 1.2.3).

Table 5.1.1B Chromosome position of the studied P-element lines

P-element	Rescue	PCR primers /Sequencing primer	Gene of Insertion*	Locus	Accession no/ Position**
H14 [Ry1] >	5'	LACF+LACS LACS	CG17508+	2R, 41E6	AE003786 7098-7099
H14 [Ry2] >	5'	LACF+LACS LACS	LD28616-***	2R, 42E4	AE003790 180054-180055
A27 [Ry2] >	5'	LACF+LACS LACS	LD28616-	2R, 42E4	AE003790 180054-180055
A35 [Ry2] >	5'	LACF+LACS LACS	LD28616-	2R, 42E4	AE003790 180054-180055
BL-11412 [PZ] >	5'	n/a	CG17800- *Dscam*	2R; 43B1-B3	AE003841*** 111005-111006 AQ034165
k16128 [lacW] >	n/a	n/a	CG1708-*cos*	2R; 43B3	AE003841***
BL-11156 [lacW] >	5'	n/a	CG1708-*cos*	2R; 43B3	AE003841*** 132448-132449 AQ034038
BL-11382 [PZ] <	3'	Pry1+ Pry4 Pry4	CG1845-	2R, 43A4	AE003841 23983-23984
k13522 [lacW] <	3'	Pry1+ Pry4 Pry4	CG2604-	3R; 82F4	AE003603 77624-77625
654 [PZ-1] <	3'	Sp3+Pry4 Pry4	CG7620+ *(l(3)87Df)*	3R, 87D12	AE003698 110346-110347
654 [PZ-2] >	5'	Sp1+SpLac2 SpLac2	CG11143- *(Inos)*	2R; 43C6-C7	AE003841 203987-203988
BL-10818 [lacW] >	5'	Sp3+Pry4 Pry4	CG2064+	2R; 43E11	AE003840 143985-143986 AQ034113#
BL-11342 [PZ] >	3'	Pry1+ Pry4 Pry4	CG11140- *Dhap*	2R, 43D1	AE003841 218999-219000
05467 [PZ] >	3'	Pry1+ Pry4 Pry4	CG11140- *Dhap*	2R; 43D1	AE003841 218999-219000
4403 [lacW] <	3'	Pry1+ Pry4 Pry4	CG2064+	2R; 43E11	AE003840 143978-143979
10467 [lacW] <	5'	Sp1+SpLac2 SpLac2	CG2140+ *(Cyt-b5)*	2R, 43E1	AE003840 16883-16884

* A "+" or "-" indicates whether the gene is on the + or - DNA strand.

**The P-element is inserted between the base pairs indicated.

***LD28616 is an EST for gene that is located between CG15242 and CG3567 (Cyp6u1)

Sequenced by Spradling et al. (1999)

< or > shows the orientation of the P-element

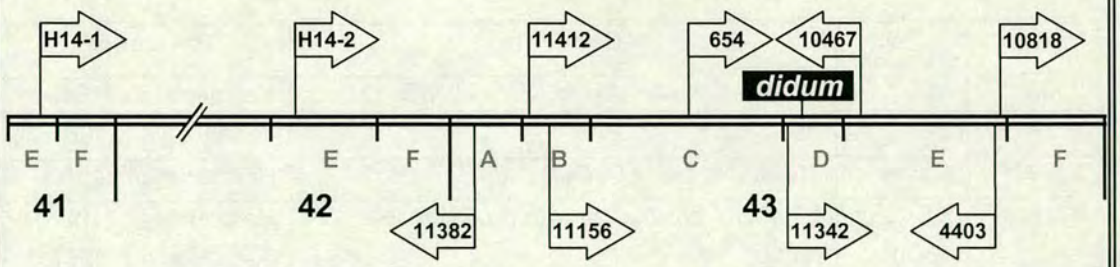


Figure 5.1.1E Position and orientation of the analysed P-element lines

The chromosome position of the P-element was determined by comparing the isolated flanking sequence to *Drosophila* genomic sequence. The genes in which the studied P-elements are inserted are listed in Table 5.1.1B. The genetic position of *didum* is shown in bold. Line H14 has two P-elements designated as 1 and 2. The H14-2 P-element is equivalent to the P-element in A27 and A35 lines. Line 654 also contains two P-elements. Only one of them is shown on the figure, the second one being located on a different chromosome (3R). Line K13522 is not shown because the only P-element it contains is inserted on the 3rd chromosome (Table 5.1.1B).

5.1.2 Search for deficiency lines

To elucidate the function and the molecular interactions of *Drosophila* Myosin V it was necessary to isolate a line(s) with a mutant phenotype. A database search found no mutant strains for *didum*, however 19 potential deficiency lines were revealed (Table 5.1.2). Deletions in the deficiency lines may cause alterations in the size of restriction fragments. The latter can be detected by Southern analysis where they appear as bands additional to the bands produced in wild-type DNA restriction digests.

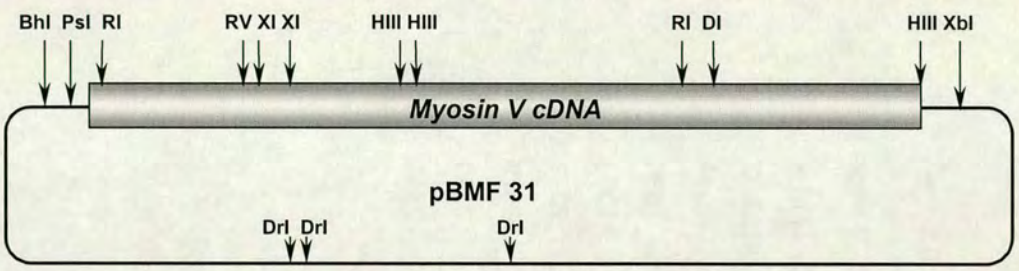
Table 5.1.2 Deficiency lines used in the screen for *didum* mutants

Stock	Map Position	Genotype	Reference
2395	NCX5	*Df(2R)NCX5 cn^1 bw^{1}sp^1 / In (2LR)O, Cy dp^{lv1}pr^1 cn^2*	Heitzler et al. (1993) John Roote (1995)
42896G	43Cc^1	*1(2)43Cc^1 bw^D / CyO, Cy^1 dp^{lv1}pr^1 cn^2*	Heitzler et al. (1993)
2398	NCX6 or 43Cb^1	*l(2)43Cb^1 cn^1 bw^1 sp^1 / In (2LR)O, Cy dp^{lv1}pr^1 cn^2*	Heitzler et al. (1993) John Roote (1995)
2399	NCX7 or 43Cc^{ncx7}	*l(2)43Cc^4 cn^1 bw^1 sp^1 / In (2LR)O, Cy dp^{lv1}pr^1 cn^2*	Heitzler et al. (1993) John Roote (1995)
4120	43Cb^1	*l(2)43Cb^1 cn^1 bw^1 sp^1 / CyO*	Heitzler et al. (1993)
4121	43Cc^4	*l(2)43Cc^4 cn^1 bw^1 sp^1 / CyO*	Heitzler et al. (1993)
4124	43Da^1	*l(2)43Da^1 / CyO*	Heitzler et al. (1993)
4125	43Db^1	*l(2)43Db^1 bw^D, Dp(?;2)bw^D / CyO*	Heitzler et al. (1993)
4126	dpa^1	*dpa^1 bw^D, Dp(?;2)bw^D / CyO*	Heitzler et al. (1993)
4129	43Ea^1	*l(2)43Ea^1 bw^D, Dp(?;2)bw^D / CyO*	Heitzler et al. (1993)
42896E	43Ca, or dpa^1	*dpa^1 bw^D / CyO, Cy^1 dp^{lv1}pr^1 cn^2 synonym l(2)43Ca^1*	Heitzler et al. (1993) John Roote (1995)
42896I	43Db^1	*1(2)43Db^1 bw^D / CyO, Cy^1 dp^{lv1}pr^1 cn^2*	Heitzler et al. (1993)
42896J	43Ea^1	*1(2)43Ea^1 bw^D / CyO, Cy^1 dp^{lv1}pr^1 cn^2*	Heitzler et al. (1993)
45987	Drl^{rv17}	*Df(2R)Drl^{rv17} / CyO, Cy^1 dp^{lv1}pr^1 cn^{2}synonym Df(2R)Drl^{Rv17}*	Heitzler et al. (1993)
45992	Drl^{rv22}	*Df(2R)Drl^{rv22} / CyO, Cy^1 dp^{lv1}pr^1 cn^2 synonym Df(2R)Drl^{Rv22}*	Heitzler et al. (1993)
2379	T(2;3)H36	*T(2;3) H36, l(2)43Ca^{36}/In(3LR)TM3, y$^+$ ri p^p sep Sb bx^{34} e^s Ser*	John Roote (1995)
2579	cos-2	*Df(2R)cos-2 cn bw sp/In(2LR)O, Cy dp^{lv1} pr cn^2*	Heitzler et al. (1993)
2580	cos-3	*Df(2R)cos-3 cn bw sp/In(2LR)O, Cy dp^{lv1} pr cn^2*	Heitzler et al. (1993)
2614	Drl^{rv28}	*Df(2R)Drl^{rv28}/In(2LR)O, Cy dp^{lv1} pr cn^2*	Heitzler et al. (1993)

A series of genomic Southerns were carried out using the potential deficiency lines. At this time we did not have the full length cDNA clones (M1-M4 clones) isolated and we used pBMF31 to generate DNA probes to be used in the Southern analyses (Fig. 5.1.2A).

Genomic DNA was prepared from heterozygous flies, strains 42896G, 2398, 2399, 2395, 2379, 2579, 2580, 2614 (the other strains shown in the table, 5.1.2 were not available for experiments at this point). Oregon R (wild-type) genomic DNA was used as a control. 10μg of genomic DNA was digested with *EcoRI, XhoI* and *PstI* restriction endonucleases and separated on 1.0 % agarose gel overnight (see materials and methods, 2.5.13 chapter). Plasmid DNA from pBMF31 was labelled

radioactively with ^{32}P and used in the subsequent hybridisations. We labelled total plasmid DNA because the labelled vector DNA from the construct hybridises with the DNA marker used in the elelectrophoretic separation of the genomic samples. Thus we were able to visualise the marker on the film. The results are shown in Fig. 5.1.2B; 5.1.2C and 5.1.2D.



Restriction map:

***EcoRI* (RI): 2.2 + 3.8**
***HindIII* (HIII): 2.0 + 4.0**
***XhoI* (XI): 0.3 + 2.3 + 3.5**
***BamHI* (BhI): 6.0**
***DraI* (DI): 17bp + 0.7 + 1.2 + 4.1**
***Xba* (XbI): 6.0**
***EcoRV* (RV): 6.0**
***EcoRI+XhoI*: 0.2 + 0.5 + 0.8 + 1.5 + 3.0**
***EcoRI+HindIII*: 0.8 + 1.0 + 1.2 + 3.0**
***EcoRI+BamHI*: 2.2 + 3.8**
***EcoRI+DraI*: 17bp + 0.15 + 0.7 + 1.2 + 1.9 + 2.2**

Figure 5.1.2A Restriction map of pBMF31

The restriction map was produced using a number of different enzymes conventionally included in restriction digests. BMF31 was the most complete cDNA available at the start of this work (MacIver et al.1998). The DNA was used to generate radioactively labelled probes to be used in Southern analysis experiments.

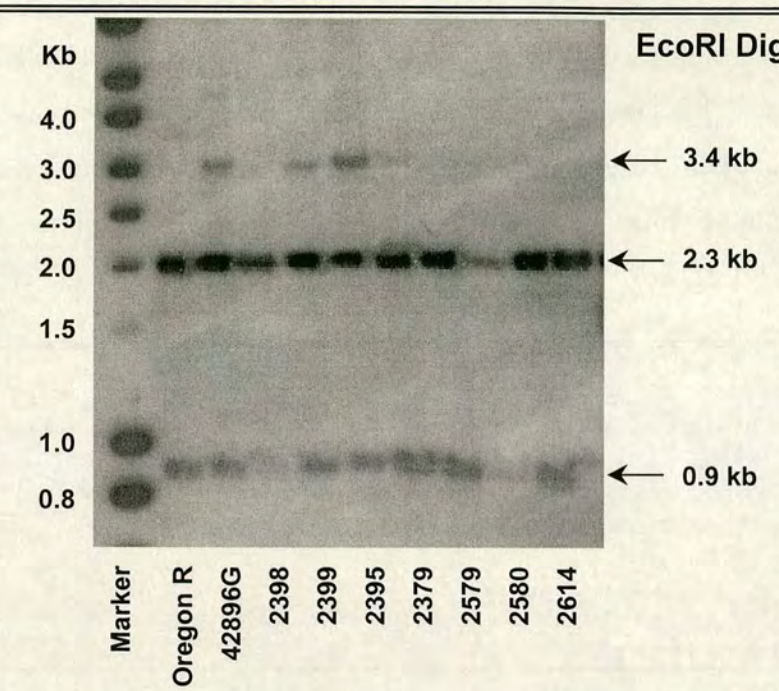


Figure 5.1.2B Deficiency lines analysis-EcoRI digest

Southern blot of genomic DNA from deficiency lines (see Table, 5.1.2). *EcoRI* digest produced 2.3 and 0.9 bands. The additional band (3.4 kb) in 42896G, 2399 and 2395 is marked with an arrow. The marker used is HyperLadder (Bioline).

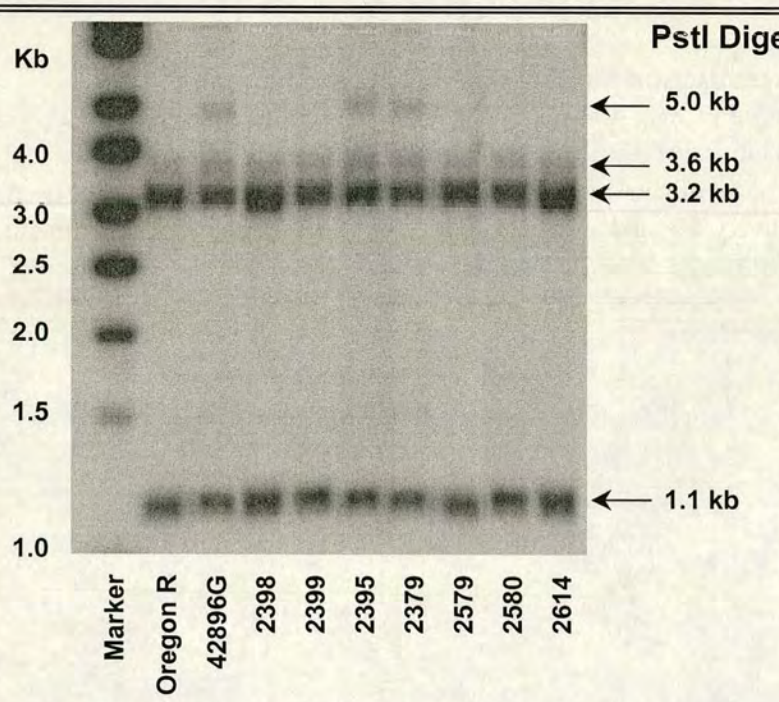


Figure 5.1.2C Deficiency lines analysis-PstI digest

Southern blot of genomic DNA from deficiency lines (see Table, 5.1.2). *PstI* digest produced 1.1, 3.2 and 3.6 kb bands. The additional band (~5.0 kb) in 42896G, 2395 and 2379 is marked with an arrow. The marker used is HyperLadder (Bioline).

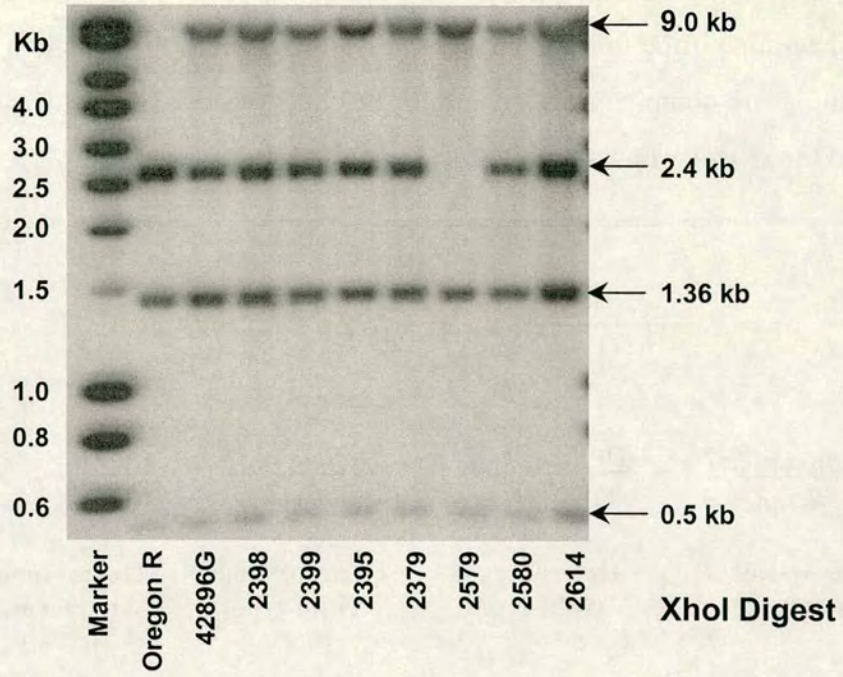


Figure 5.1.2D Deficiency lines analysis-XhoI digest

Southern blot of genomic DNA from deficiency lines (see Table, 5.1.2). *XhoI* digest produced 0.5, 1.36, 2.4 and 9.0 kb bands. No additional bands were observed in *XhoI* digest. Interestingly Oregon R (wild-type) genomic DNA lacked the 9.0 kb band and 2579 lacked the 2.4 kb band. This is probably due to polymorphism of the background strains used to generate the deficiencies. The marker used is HyperLadder (Bioline).

The Southern Blot analysis showed that four of the studied deficiency lines produce an extra band when restricted with *EcoRI* and *PstI* but not *XhoI*. 42896G, 2399, 2395 and 2379 strains were chosen for further analysis.

5.1.3 Complementation mapping

To determine if the four deficiency lines, isolated from the Southern analysis belong to the same complementation group, the lines were crossed to each other according to the scheme shown on Fig. 5.1.3A.

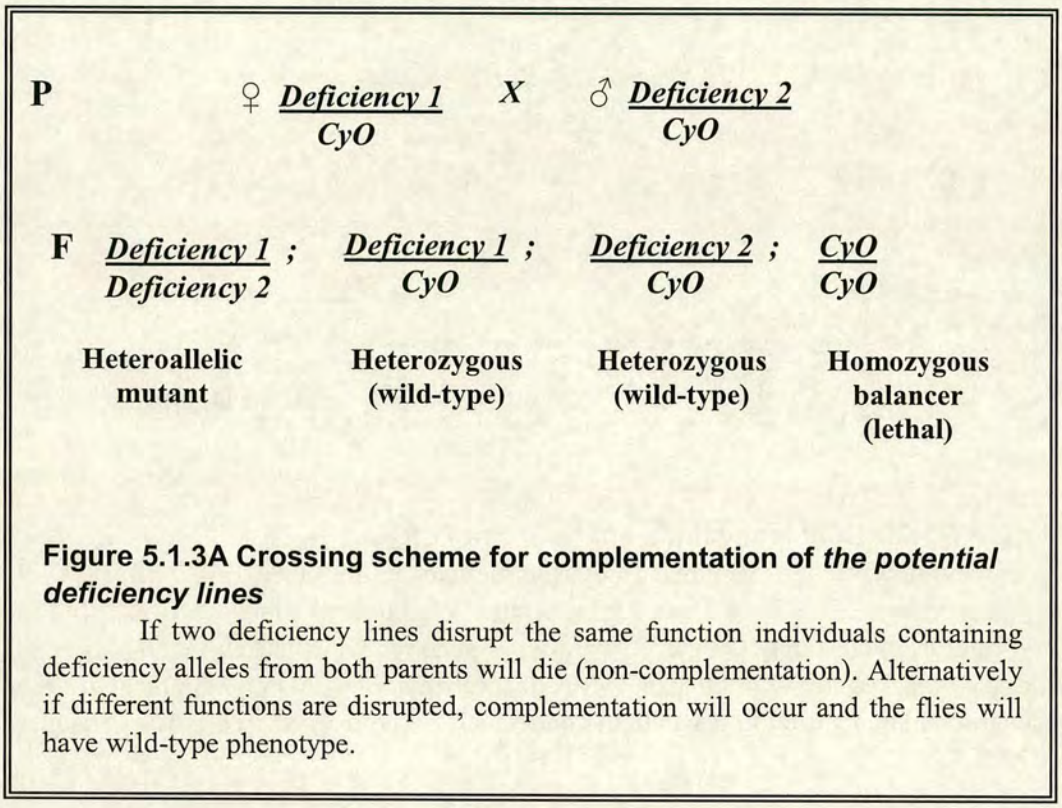


Figure 5.1.3A Crossing scheme for complementation of *the potential deficiency lines*

If two deficiency lines disrupt the same function individuals containing deficiency alleles from both parents will die (non-complementation). Alternatively if different functions are disrupted, complementation will occur and the flies will have wild-type phenotype.

Complementation analysis of 42896G, 2399, 2395 and 2379 showed that each cross produces wild-type flies. This indicated that the four lines are in different complementation groups and disrupt different functions.

During the preliminary complementation test we were able to collect a number of new deficiency lines (listed in Table 5.1.2). At the same time the Genome sequence of *Drosophila melanogaster* was released. Thus we were able to select lines that map very closely to *didum*. Lines 42896G, 2399 and 2395 but not 2379 were shown to map closely to *didum* and was analysed further.

The new deficiency lines as well as the three lines from the previous analysis were subjected to complementation crosses. For each cross 2-3 female virgin flies were crossed to 3 male flies. Approximately 80 to 100 individuals from the offspring were counted for statistically representative results. The complementation results are presented in Table 5.1.3.

Table 5.1.3 Complementation analysis of the potential deficiency lines

	2395	42896G	2398	4120	2399	4121	4124	4125	42896I	42896E	4126	4129	42896J	45987*	45992*
2395 NCX5	-	-	+	+	+	+	+	+	+	+	+	+	+	-	-
42896G Cc^{1}	-	-	+	+	+	+	+	+	+	+	+	+	+	-	-
2398 Cb^{1}	+	+	-	-	+	+	+	+	+	+	+	+	+	-	-
4120 Cb^{1}	+	+	-	-	+	+	+	+	+	+	+	+	+	-	-
2399 Cc^{4}	+	+	+	+	-	-	+	+	+	+	+	+	+	-	-
4121 Cc^{4}	+	+	+	+	-	-	+	+	+	+	+	+	+	-	-
4124 Da^{1}	+	+	+	+	+	+	-	+	+	+	+	+	+	-	-
4125 Db^{1}	+	+	+	+	+	+	+	-	-	+	+	+	+	+	-
42896I Db^{1}	+	+	+	+	+	+	+	-	-	+	+	+	+	+	-
42896E dpa^{1}	+	+	+	+	+	+	+	+	+	-	-	+	+	-	-
4126 dpa^{1}	+	+	+	+	+	+	+	+	+	-	-	+	+	-	-
4129 Ea^{1}	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+
42896J Ea^{1}	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+
45987* Drl^{rv17}	-	-	-	-	-	-	-	+	+	-	-	+	+	-	-
45992* Drl^{rv22}	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-

* These lines are large deletions in chromosome region 43Ca-43Da.

Using the data from the complementation test and cytological data from the Flybase (http://flybase.bio.indiana.edu/) we were able to built a complementation map (Fig. 5.1.3B).

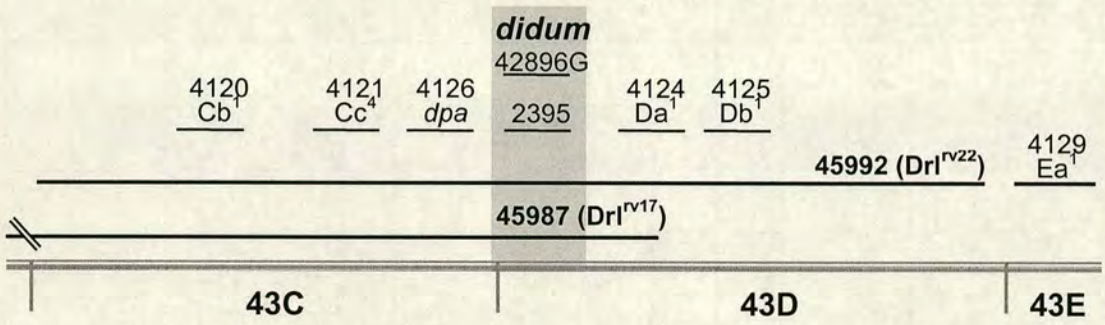


Figure 5.1.3B Complementation map

The analysed deficiency lines were placed in ten complementation groups. For lines that fall in the same complementation group only one is shown on the diagram. These lines are: 2398 group with 4120, 2399 with 4121, 42896I with 4125, 42896J with 4129, and 42896E with 4126. The relative position of *didum* (*myosin V*) on the chromosome map is shaded.

5.2 *Analysis of the isolated potential Myosin V mutants*

5.2.1 Generating GFP-balanced deficiencies

To check whether the potential 2395 and 42896G lines are mutants for myosin V we needed to isolated homozygous individuals. Since both lines are homozygous lethal we marked the deficiency alleles with GFP (green fluorescent protein) marker. Thus when self crossed the flies would produce 25% non-fluorescent individuals that are homozygous for the deficiency (Fig. 5.2.1A).

Initially we used a GFP marker where the GFP is fused to Kr (Fig. 5.2.1A). Strains 2395, 4124 (Da), 4125 (Db), Drl^{rv17} and Drl^{rv22} were crossed to Kr-GFP balancer strain. We found it impossible to generate a 4125/GFP line with the shown cross. No offspring were produced when GFP virgin females were crossed to deficiency males. To overcome this problem we performed a reciprocal cross with virgin females (deficiency) with GFP males (Fig. 5.2.1B).

We found that this fluorescent balancer is difficult to work with because it produces a weak fluorescence almost indistinguishable from the autofluorescence of the fly embryos. Thus it was difficult to separate fluorescent from non-fluorescent embryos. In our subsequent experiments we used a GFP marker (GFP-Actin) Fig.5.2.1C).

P ♀ ***L, Pin / CyO, Kr-GFP*** **X** ♂ ***Deficiency / CyO***

F_0	***L, Pin / Deficiency***	***L, Pin / CyO***	***Deficiency / CyO, Kr-GFP***	***CyO / CyO, Kr-GFP***
	Discard flies with small eyes (L) and short, thick thoracic bristles (Pin)		**Select for Curly wings (Cy)**	**Lethal**

P_1 ♀ ***Deficiency / CyO, Kr-GFP*** **X** ♂ ***Deficiency / CyO, Kr-GFP***

Keep the progeny as a stock:

F_1	***Deficiency / Deficiency***	***CyO, Kr-GFP / CyO, Kr-GFP***	***Deficiency / CyO, Kr-GFP***
	Embryos, larvae and adults do not show fluorescence	**Embryonic lethal**	**Embryos, larvae and adults show fluorescence**

Figure 5.2.1A Cross for generating GFP balanced deficiency lines

The original CyO (wing) marker was crossed out and substituted for a fluorescent GFP marker, Kr-GFP. The generated flies with a deficiency over Kr-GFP were self crossed and established as stock.

P ♀ *Deficiency / CyO* **X** ♂ *L, Pin / CyO, Kr-GFP*

F_0 *L, Pin / Deficiency* | *L, Pin / CyO* | *Deficiency / CyO, Kr-GFP* | *CyO / CyO, Kr-GFP*

Discard flies with small eyes (L) and short, thick thoracic bristles (Pin) | **Select for Curly wings (Cy)** | **Lethal**

P_1 ♀ *Deficiency / CyO, Kr-GFP* **X** ♂ *Deficiency / CyO, Kr-GFP*

Keep the progeny as a stock:

F_1 *Deficiency / Deficiency* | *CyO, Kr-GFP / CyO, Kr-GFP* | *Deficiency / CyO, Kr-GFP*

Embryos, larvae and adults do not show fluorescence | **Lethal** | **Embryos, larvae and adults show fluorescence**

Figure 5.2.1B Cross for generating GFP balanced deficiency lines

A reciprocal cross for generating 4195/GFP line. The original CyO (wing) marker was crossed out and substituted for a fluorescent GFP marker, Kr-GFP. The generated deficiency over Kr-GFP were self crossed and established as stock.

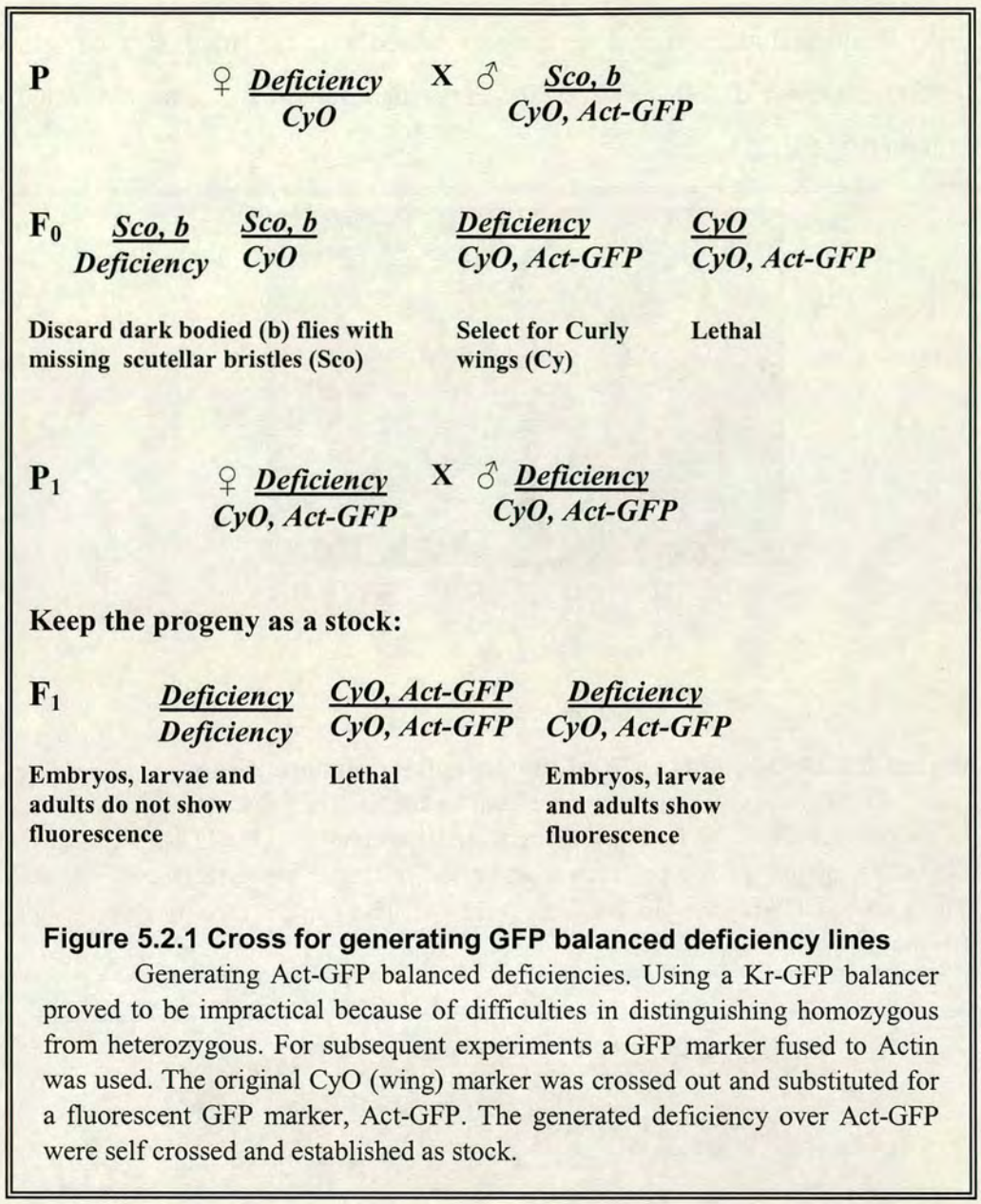


Figure 5.2.1 Cross for generating GFP balanced deficiency lines

Generating Act-GFP balanced deficiencies. Using a Kr-GFP balancer proved to be impractical because of difficulties in distinguishing homozygous from heterozygous. For subsequent experiments a GFP marker fused to Actin was used. The original CyO (wing) marker was crossed out and substituted for a fluorescent GFP marker, Act-GFP. The generated deficiency over Act-GFP were self crossed and established as stock.

5.2.2 Molecular analysis of the deficiency lines

Genomic DNA was isolated from deficiency lines 2395, 4124 (DA), 4125 (Db) and 42896G. Genomic DNA from Oregon R (wild-type) was used as a control. The DNA was amplified in a PCR reaction, using 5'UTR and MR3 primers (Fig. 5.2.2A). The PCR amplified approximately 9.0 kb of genomic DNA that covered completely the *myosin V* gene. The products were restricted with *XhoI, XbaI* and *EcoRI* endonucleases in single and double digests (Fig.5.2.2B). We expected that when digested the DNA from heterozygous flies would produce extra band(s) on agarose gel. This is possible only when the aberration is a significant duplication or

deletion. Point mutations would remain undetected with this method. Generating a restriction map would help us to identify in which part of the gene is located the aberration (Fig.5.2.2C).

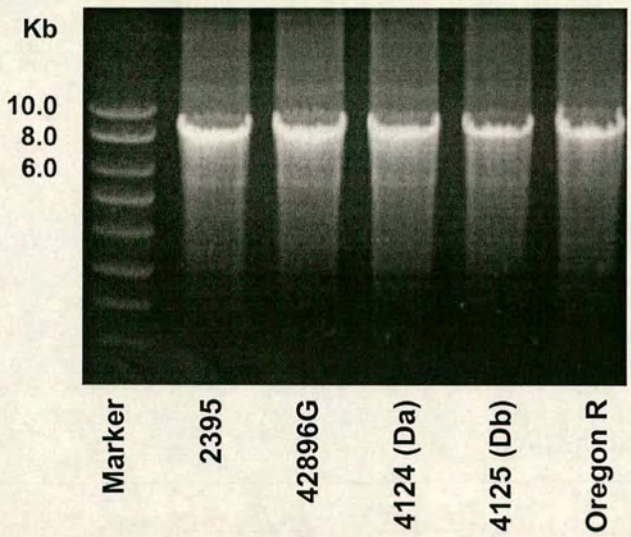


Figure 5.2.2A Genomic PCR of the potential deficiency lines

5'UTR and MR3 primers were used to amplify the 9.0 kb genomic sequence of *didum* (*myosin V)*. To optimise the amplification conditions the PCRs were carried out with a mixture of *Taq* polymerase and proof-reading *Pfu* polymerase in 3:1 ratio. The elongation time was 15 minutes at 72 °C. The marker used is HyperLadder (Bioline).

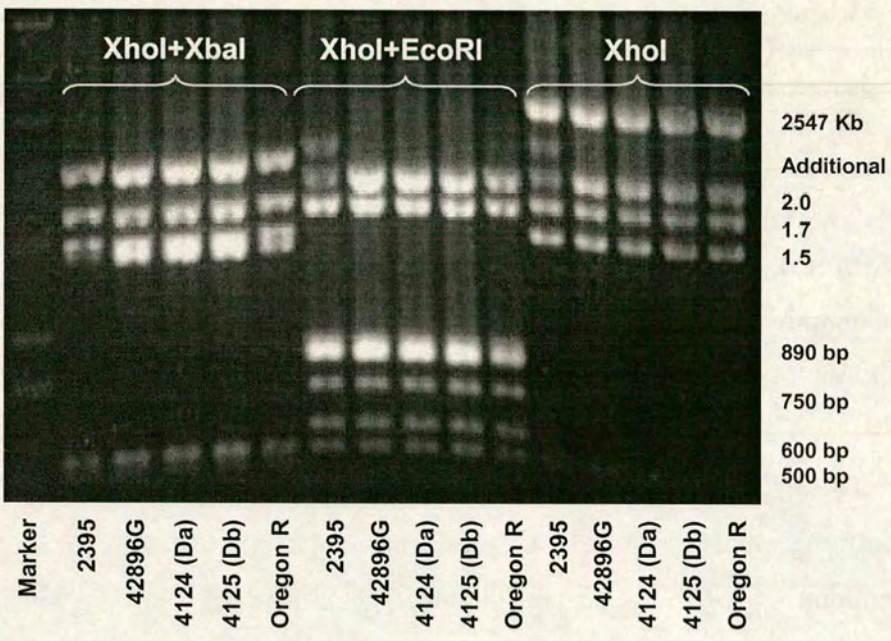


Figure 5.2.2B Restriction digest of 5UTR+MR3 PCR product

The restriction analysis revealed an additional band in line 2395 digest. To map precisely the aberration site further digests were needed. These are shown on Fig. 5.2.2C.

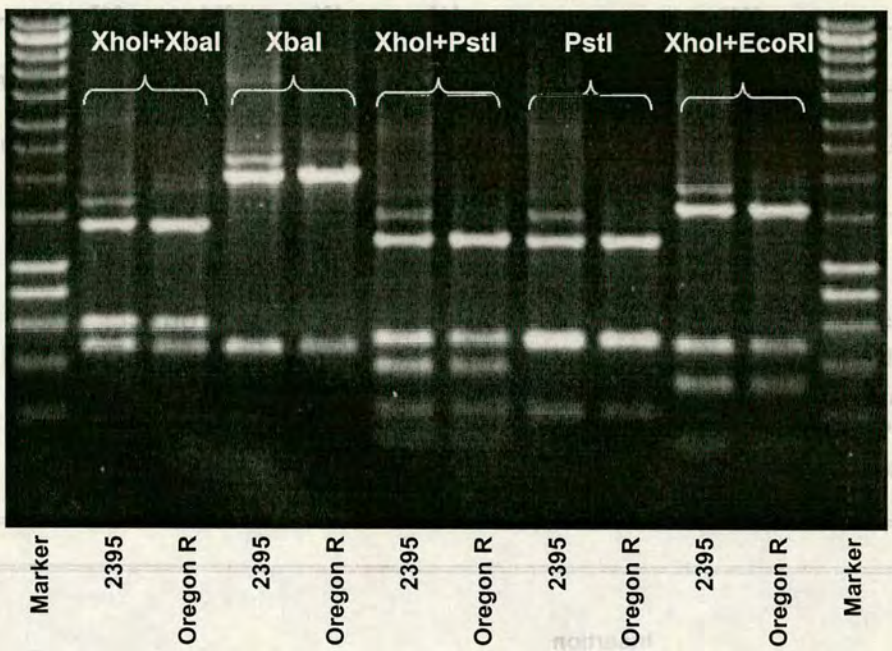


Figure 5.2.2C Restriction digest of 2395 (5UTR-MR3 PCR product)

All digests were carried out with ~500ng of the 5UTR+MR3 PCR amplified product. The reaction were performed overnight to ensure completion of the reaction. Oregon R (wild-type) was used as a background control. The marker used is Hyper Ladder (Bioline).

A restriction map was built using the data from the two restriction digests (Fig5.2.2D). Analysis of the restriction products showed the presence of approximately 150-200 bp insertion in the region between *XbaI* and *PstI* restriction sites. A PCR approach was chosen to map down the exact location of the extra DNA. 5UTR, STM2, did5, did10 and BSM1 primers were used to amplify the region between XbaI and PstI restriction sites. Since the template in that PCR was from heterozygous flies we expected to obtained two amplification products; one wild-type sequence and the mutant sequence containing the additional DNA. The PCR did produce two amplification products (Fig. 5.2.2E). Interestingly the mutant sequence was always much weaker than the wild-type.

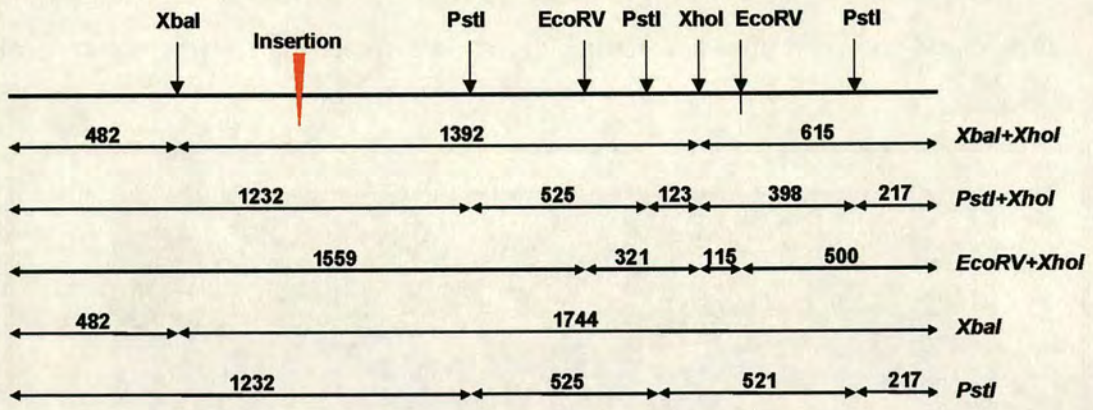


Figure 5.2.2D Restriction map of 5UTR+MR3 amplified product

Restriction digest was carried out with selected restriction endonucleases. Comparison of the restriction products generated from wild-type sequence with these from 2395 sequence showed the presence of additional , ~200 bp longer band in the mutant sequence. The insertion site is shown in red.

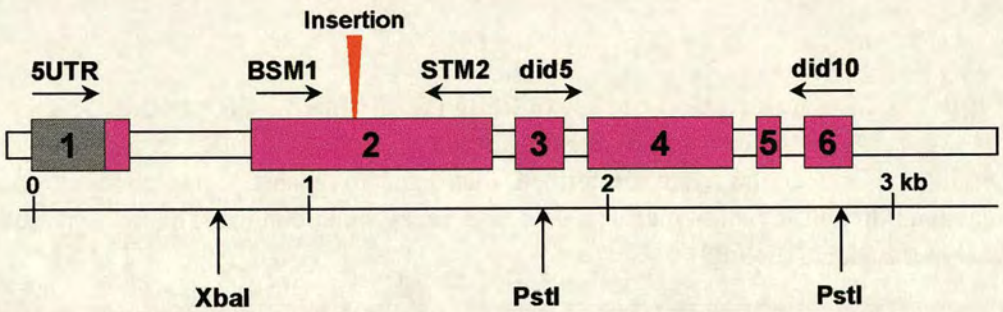


Figure 5.2.2E PCR analysis of XbaI-PstI region

PCR was carried out with the following primer pairs: 5UTR+STM2, 5UTR+did10, BSM1+STM2, BSM1+did10, did5+did10. For the subsequent experiments only 5UTR+did10 and BSM1+STM2 amplification products were used.

The products of 5UTR +did10 and BSM1+STM2 were purified from agarose gel and blunt-end cloned in TOPO vector (see materials and methods). To check for correct cloning the recombinant TOPO constructs were digested with *EcoRI*. Since there are two *EcoRI* site in the TOPO linker sequence, one on each site of the cloned

sequence a restriction would produce two products: the vector sequence itself and the full length cloned sequence (Fig. 5.2.2F).

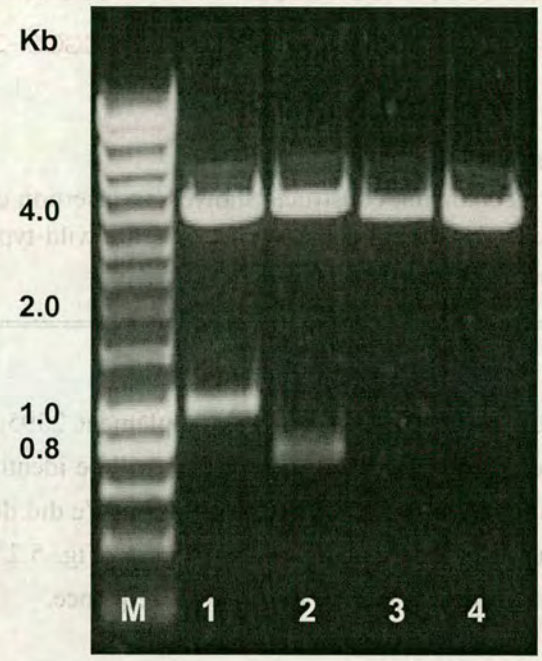


Figure 5.2.2F Restriction analysis of the subcloned STM2+BSM1 and 5UTR+did10 sequences

STM2+BSM1 products (1kb for the mutant and 0.8 kb wild-type sequence) were cloned successfully in TOPO vector. Unfortunately none of the 5UTR+did10 products were cloned into the vector. The 4.0 kb band is the TOPO vector. The marker used is HyperLadder (Bioline).

The recombinant constructs containing the STM2+BSM1 amplified products were named as follows: pMSH, containing the mutant fragment and pWTSH containing with the wild-type sequence. Both constructs were subjected to sequencing. It was extremely hard to sequence the mutant sequence because it proved to be a repeat sequence (Fig. 5.2.2G). Approximately 200 bp of CT-rich sequence was inserted between bases 635-636 (TTGCAC-AATCTG) of *myosin V* second exon. The complementary sequence of the insert is the well known AGAG sequence found in gene enhancers.

Wild-type sequence: 5'-AATCTGCGTGTCCG-3'

Mutant sequence: 5'-GCCGGGGGTGTTGCACCTCTCCTCTCCTCTCCTCT
CCTCTCCTCTCCTCCCCT(CTCCT)$_{x150}$CTTCTGCGAGCGCC-3'

Figure 5.2.2G Mutant sequence found in strain 2395

Analysis of pMSH and pWTSH recombinant constructs showed an insertion of approximately 200 bp CT-reach sequence in the mutant construct. Instead the wild-type sequence (given in blue) the insertion (shown in red) was identified.

The complementation test showed that line 42896G fail to complement 2395. We analysed 42896G by additional restriction digests to see if any extra bands will be identified. The presence of such band could help us to map the mutation site in 42896G. We did detect an extra band of approximately 750 bp in the products of XhoI+DraI digest (Fig. 5.2.2H). Unfortunately time limitation did not allows to further analyse the 42896G sequence.

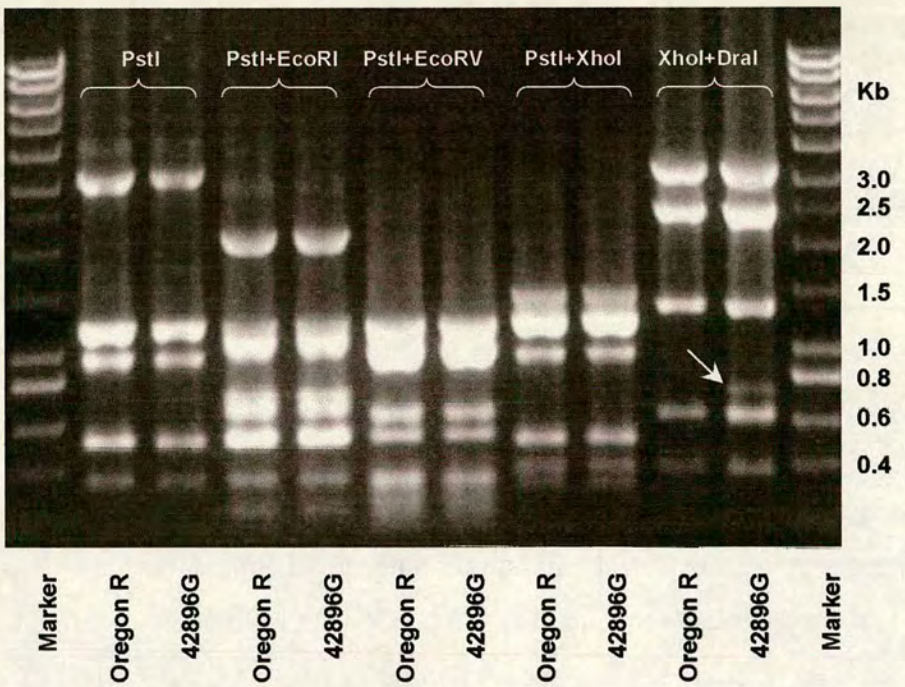


Figure 5.2.2H Restriction analysis of 42896G strain

Additional band is detected in XhoI+DraI digest (white arrow). Oregon R was used as a control. The marker used is HyperLadder (Bioline).

5.1.1 Expression analysis of the isolated mutant lines

To show that the mutation in 2395 and 42896G interfere with the normal function of *myosin V* we analysed the protein expression of *myosin V* in these lines. Homozygous embryos were collected from the GFP balanced 2395 and 42896G strains (non-fluorescent individuals) Fig. 5.2.3A). Protein extracts were prepared from the embryos and the proteins separated on PAGE. Immunodetection with the purified Myosin V antibody was carried out (Fig. 5.2.3B).

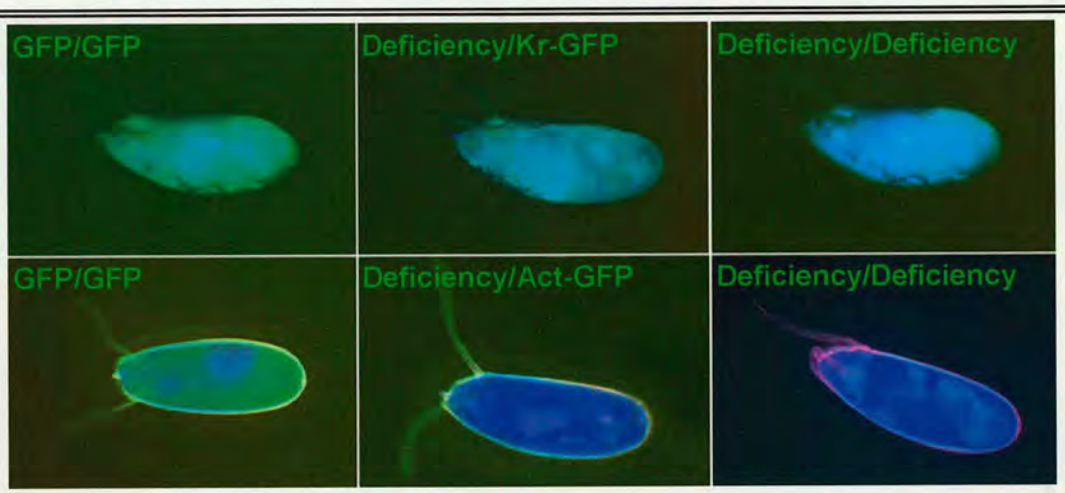


Figure 5.2.3A Distinguishing between homozygous and heterozygous GFP embryos

Initial experiments generated deficiencies balanced over Kr-GFP. As can be seen from the picture the GFP expression is weak and homozygous deficiency embryos (lack GFP balancer) can not be isolated accurately. Later experiments were performed with GFP balancer where the GFP is fused to Actin.

The Western analysis showed that both lines, 2395 and 42896G do not express Myosin V or the expression levels are reduced to levels undetectable by Western blot. As a positive control we used protein extract from wild-type Oregon R embryos. These produced a well defined band of approximately 200 kDs, the expected size for 1815 amino-acid protein. As a loading control we used specific antibodies to detect non-muscle Myosin II expression (228 kDa). The expression of Myosin II proved to be unaffected in all studied lines. These results indicate that the mutations in 2395 and 42896G disrupt the function of *Myosin V*.

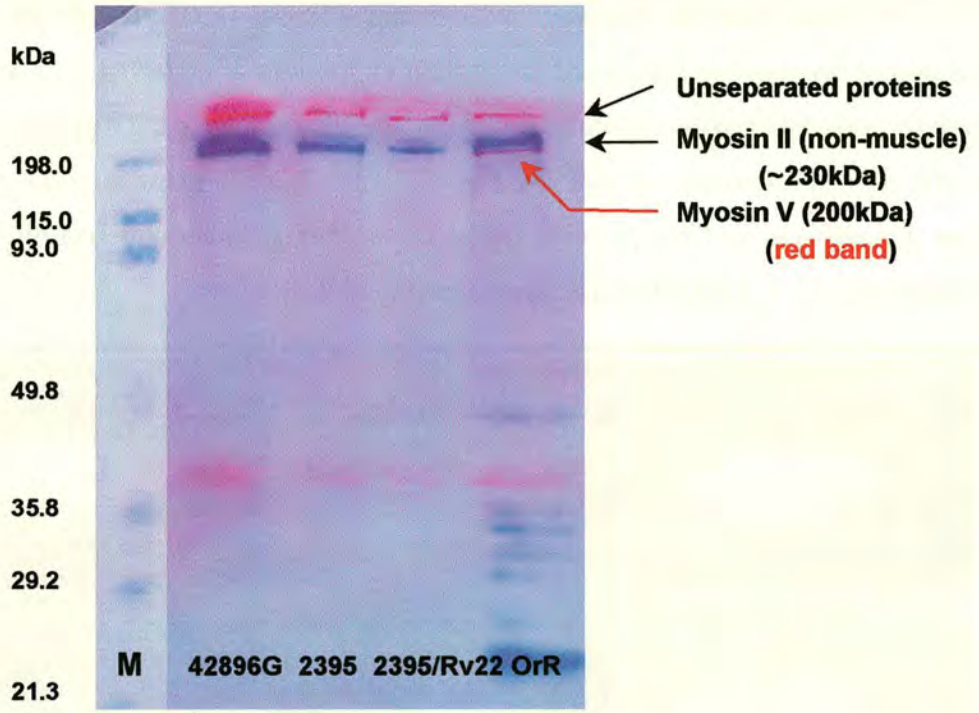


Figure 5.2.3 Western analysis of the myosin V expression in the mutant strains

Protein extracts from homozygous 42896G and 2395 larvae were analysed by immunodetection with the specific Myosin V antibodies. Oregon R embryos were used as a positive control for Myosin V wild-type expression. Protein from heteroallelic 2395/Drl^{rv22} embryos was used as a negative control. The latter would show reduced levels of Myosin V expression even if mutation in 2395 is not affecting this expression.

5.1.2 Analysis of the mutant phenotype

From the complementation test we knew that both lines, 2395 and 4289 are homozygous lethal. Even when grown in large population no escaper flies were observed. To determine the lethal stage we collected 2395/GFP and 42896G/GFP embryos. These were studied on a fluorescent microscope and photographed (Fig. 5.2.4A). All embryos, homozygous and heterozygous, were normal and hatched without problems.

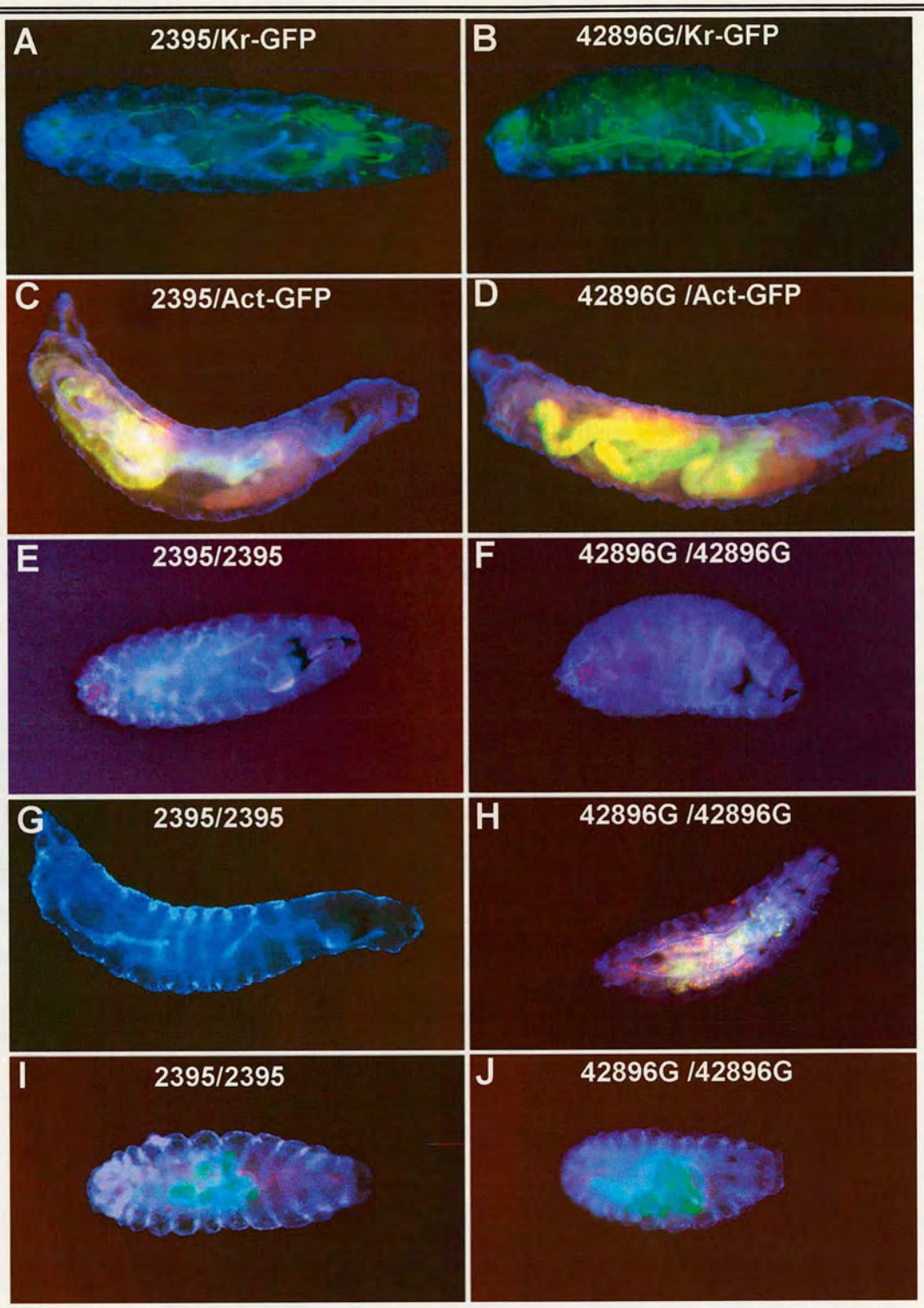


Figure 5.2.4A Mutant phenotype associated with the disruption of Myosin V function

A, B: Initial experiments with Kr-GFP balanced flies.

C, D: Heterozygous larvae from 2395 and 42896G. These larvae develop normally.

E, F, G: Homozygous mutant phenotype (no GFP expression). 2395 die by the end of first instar stage; 42896G larvae die shortly after hatching.

H, I, J: Homozygotes mutant larvae that show fluorescence due to feeding on heterozygous (GFP expressing) larvae .

We observed that approximately 96% of the heterozygous larvae from both lines proceeded through normal development. The remaining 4% showed abnormal development dying during the first instar period. Approximately 74% of the homozygous 2395/2395 larvae appeared normal until the end of the first instar stage when they died in a short period of time. The rest, 26%, had shorter and rounder bodies and showed abnormal behaviour (slow and indecisive movements). They also died by the end of the first larval instar. Occasionally we found 2395/2395 larvae that move backwards. These were examined closely but they showed normal segmentation and intact epithelial structures such as denticles and bristles. Similar phenotype was observed in homozygous 42896G individuals. More than 95% of the 42896G embryos acquired rounded bodies shortly after hatching. All individuals with 42896G mutant phenotype died before the end of first instar stage. Most of the homozygous 42896G/42896G larvae were immobile with certain individuals showing slow and uncoordinated, short distance movements.

To look for any cuticle abnormalities we mounted homozygous and heterozygous larvae in Hoyer's medium (see materials and methods). These were examined on a light transmission microscope and photographed. We found no changes in the internal structures of the heterozygous and 2395 homozygous embryos. In 5% of the examined homozygous 4896G larvae we found healed scars on the larval cuticle which we believe are from abnormal dorsal closure during embryogenesis (Fig. 5.2.4B)

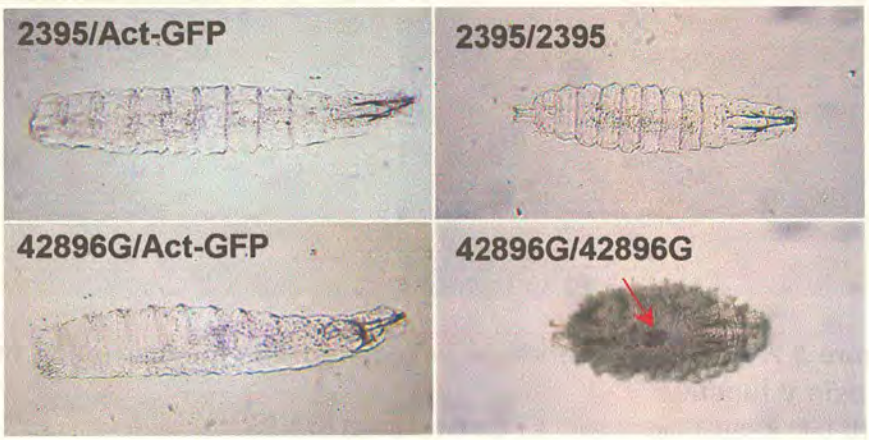


Figure 5.2.4B Mutant phenotype associated with the disruption of Myosin V function

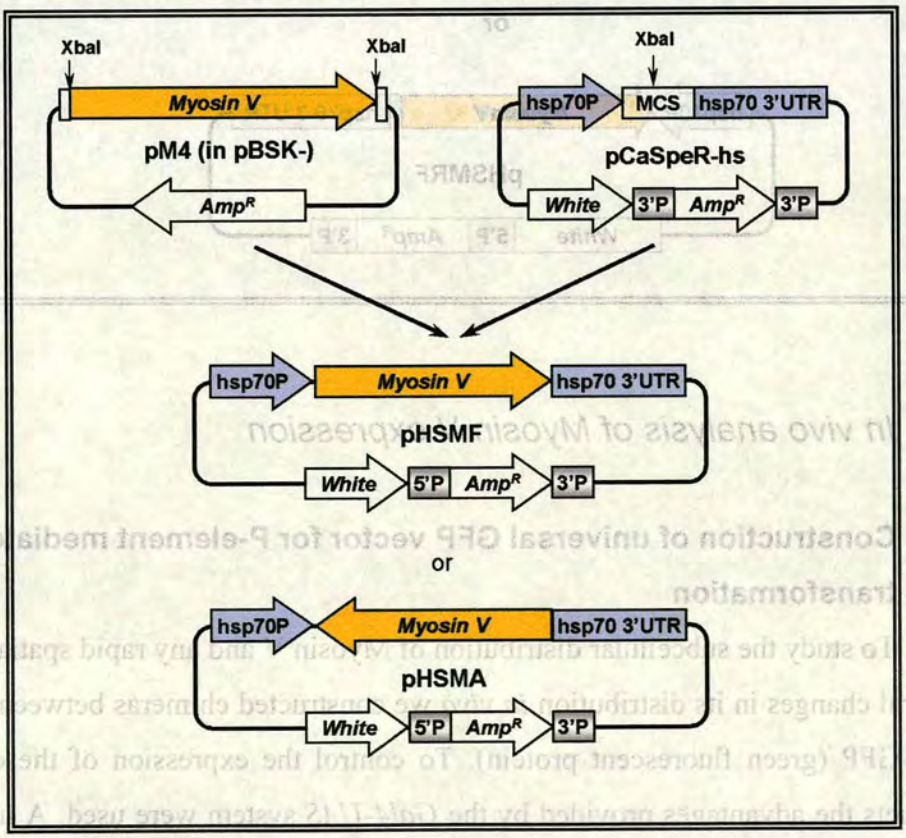
A cuticle scar was found in 5% of the 42896G homozygous embryos (red arrow).

5.2 *Targeted silencing and activation of Myosin V expression during Drosophila development*

Overexpression or suppression of the gene expression will disrupt the function of Myosin V at specific developmental stages and in specific cell types, thus allowing the function of the gene to be studied. Two different systems were chosen to carry out experiments where inducible sense or antisense Myosin V RNA is to be transformed in wild-type flies: In the *Gal4-UAS* system a tissue specific *Gal4* induces expression of a *UAS* that is fused to a sense or antisense transgene-spatial specific expression. In the second system the transgene is under the control of a heatshock promotor thus allowing us to control its expression in a temporal-specific manner. In this thesis we describe the construction of these recombinant constructs. Due to time limitations we were not able to transform and study their effect in flies.

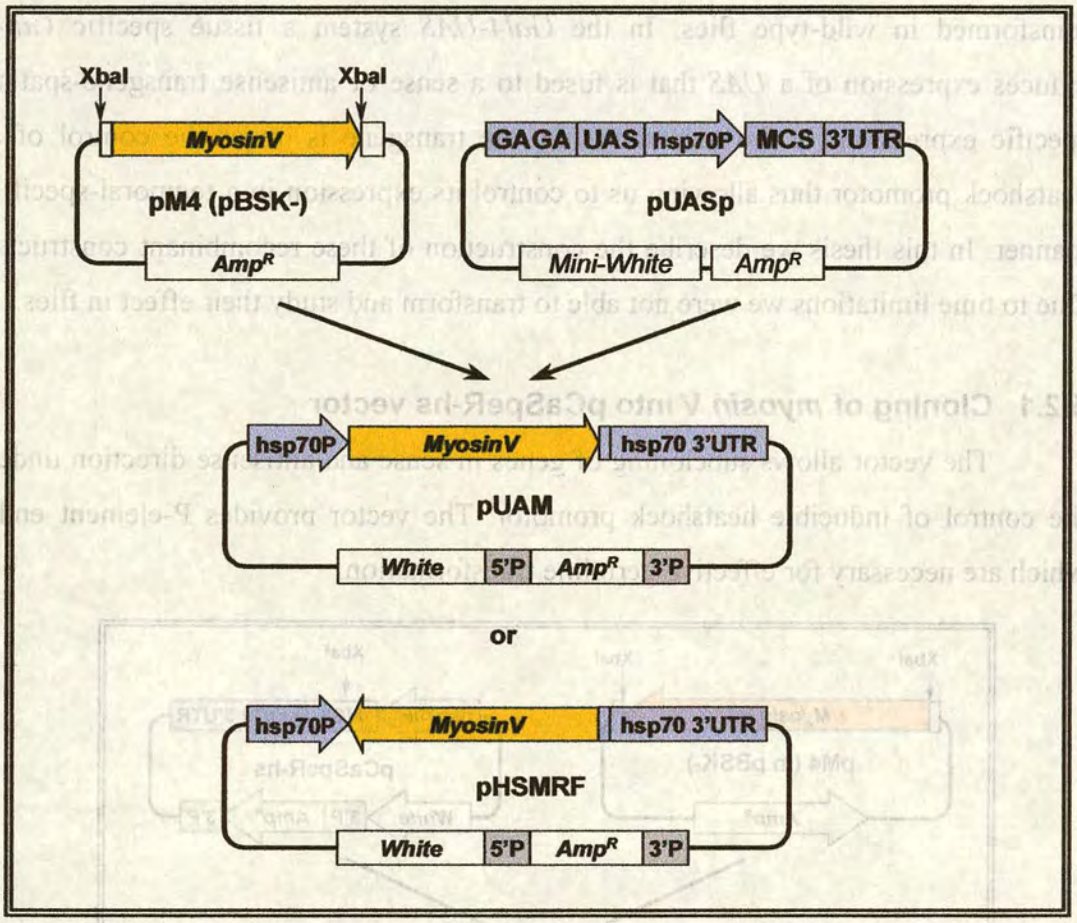
5.2.1 Cloning of *myosin V* into pCaSpeR-hs vector

The vector allows subcloning of genes in sense and antisense direction under the control of inducible heatshock promotor. The vector provides P-element ends which are necessary for effective germline transformation.



5.2.2 Cloning of *myosin V* into pUASp vector

The vector allows subcloning of genes in sense and antisense direction under the control of *UAS* (upstream activating sequence), *(Gal4* activated) driven promotor. The vector provides P-element ends necessary for germline transformation. We chose this vector because it is expressed in the germline cells (Rorth 1998), and *myosin V* is also expressed in these cells during oogenesis.

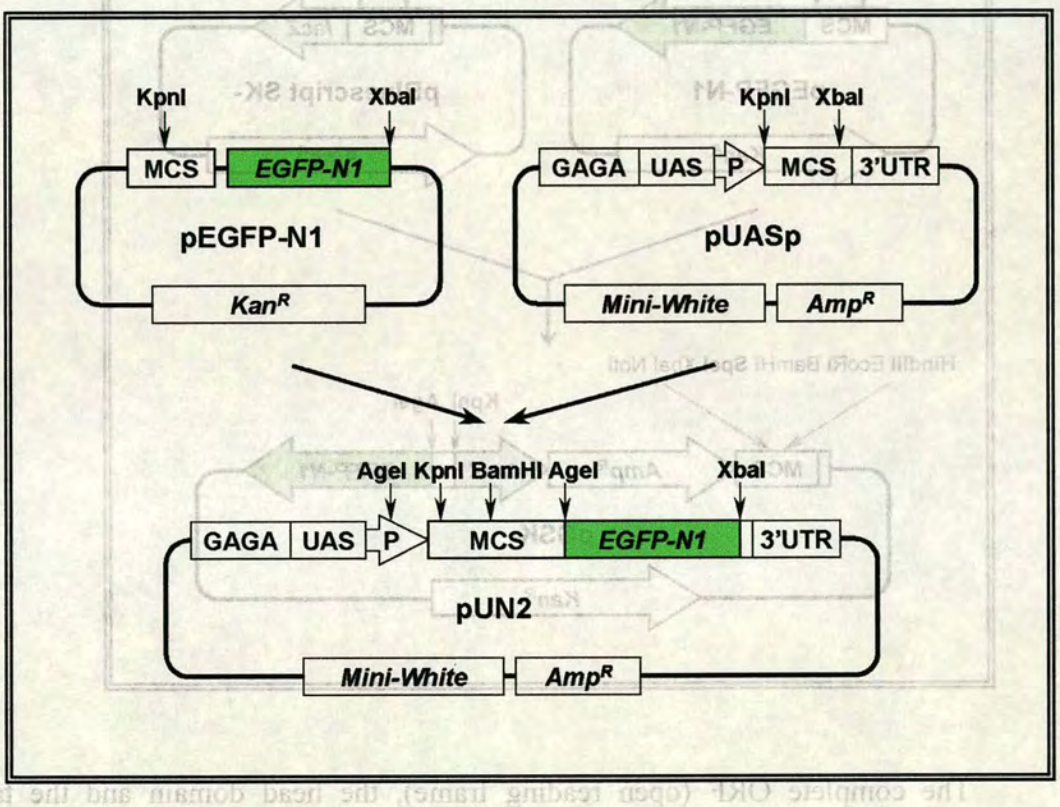


5.3 *In vivo analysis of Myosin V expression*

5.3.1 Construction of universal GFP vector for P-element mediated transformation

To study the subcellular distribution of Myosin V and any rapid spatial and/or temporal changes in its distribution *in vivo* we constructed chimeras between *myosin V* and GFP (green fluorescent protein). To control the expression of these fusion constructs the advantages provided by the *Gal4-UAS* system were used. A universal

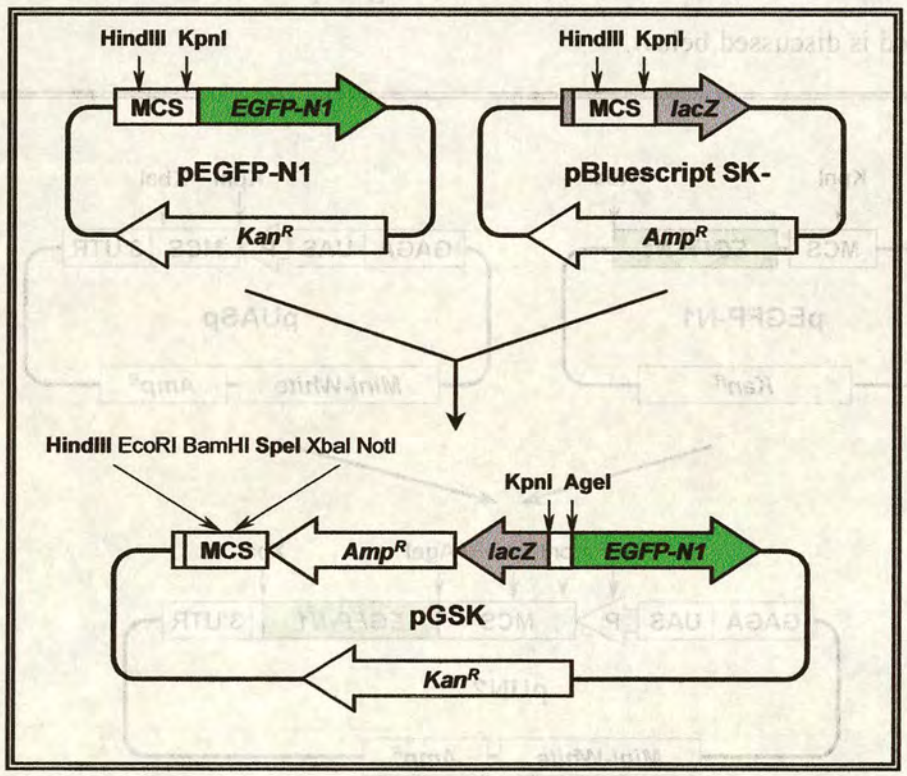
vector (pUN2), a combination of pUASp and GFPN1 vectors was constructed. Our plans were to clone different parts (domains) from myosin V in pUN2. Expressing the complete *myosin V-GFP* sequence will gives us information about the distribution of the wild-type Myosin V. Expressing only the head domain (*head-GFP*) will serve as a control (the head contains the actin binding site). Expressing the tail (*tail-GFP*) will help us to understanding the function of Myosin V, and find out what is the cargo to which the tail binds. Unfortunately the *myosin V* cDNA contained all the restriction sites need for cloning in pUN2. An alternative cloning approached is discussed below.



5.3.2 Cloning of *myosin V-GFP* sequence in pUASp

To create a fusion between myosin V and pUASp the construction of an intermediate construct was needed (Fig. 5.4.2A). For this reason we combined pEGFP-N1 with pBluescript to make the pGSK vector. Thus we could choose from wide range of restriction sites for the subsequent cloning of the *myosin V* domains.

Figure 5.4.2A Construction of pGSK vector

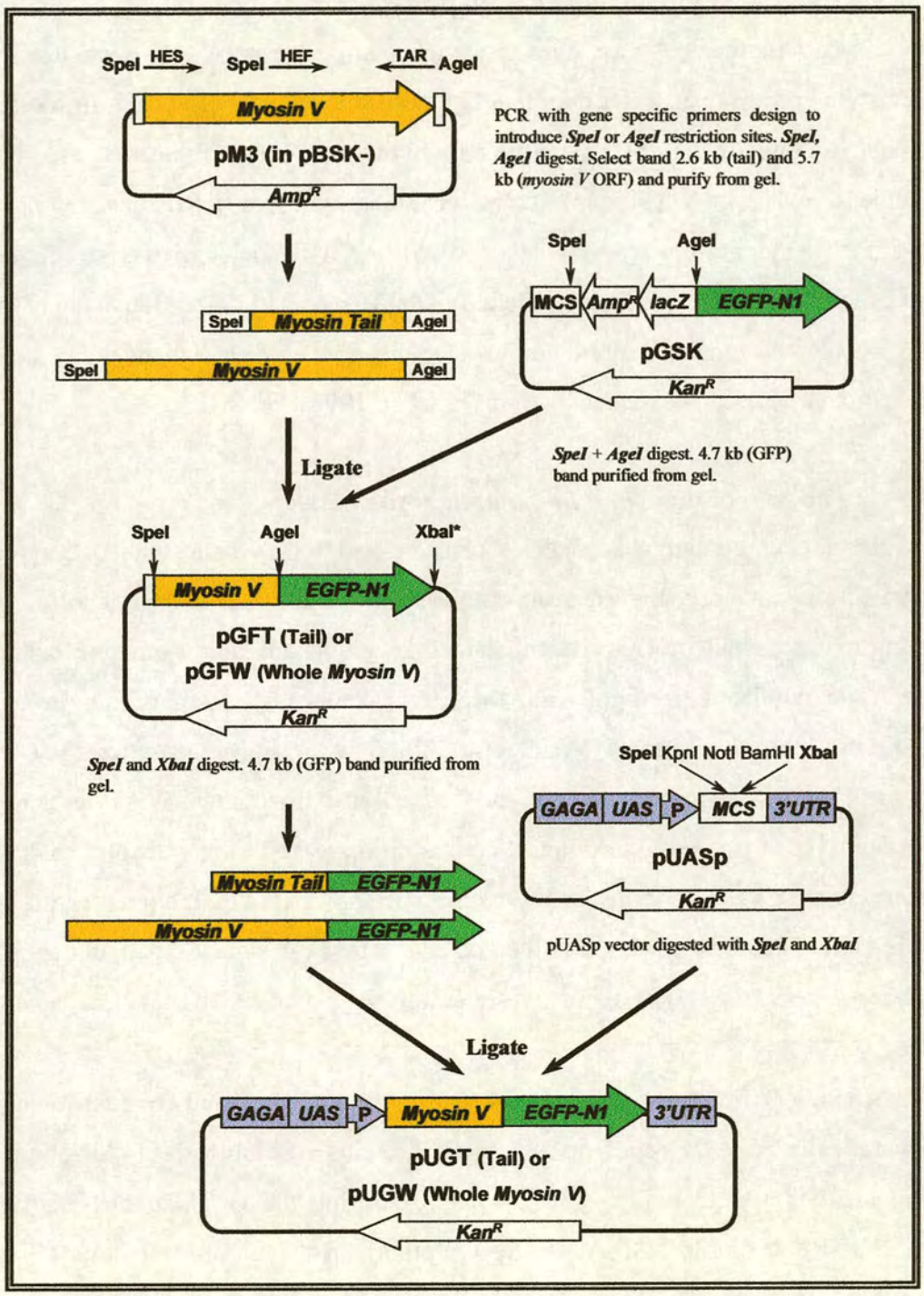


The complete ORF (open reading frame), the head domain and the tail domain of *myosin V* were amplified by PCR using specific primer which introduce *SpeI* and *AgeI* restriction site at the ends of the amplified products (see Fig. 5.4.2B for further details). The latter were digested with *AgeI+SpeI* and ligated to pGSK vector. Because of time limitations only the constructs containing the total *myosin V* cDNA and the one with the tail domain were further processed.

The *Myosin V-GFP* (total and tail domain constructs) fusions were excised from pGFT and pGFW with *SpeI+XbaI* restriction digests and cloned into pUASp to

generate pUGT and pUGW vectors. The latter could be introduced to flies using P-element mediated transformation.

Figure 5.4.2B Construction of pUGT and pUGW vectors



5.4 Discussion

5.4.1 Functions of Myosin V in species other than *Drosophila*

5.4.1.1 The role of myosin V in mRNA transport

In yeast mother cells are able to switch mating types. A HO endonuclease induces switching by causing a double stranded break at the *MAT* locus in mother cells. In daughter cells a transcriptional inhibitor Ash1p represses the HO endonuclease. Myosin V (Myo4p) is responsible for the restricted localisation of *ASH1* mRNA and protein (Jansen et al. 1996). As a result Ash1p is selectively expressed in the daughter cells only. Deletions in *Myo4* yield no obvious phenotype but *ASH1* mRNA remain in the cytoplasm of the mother cells and no long range directed movements can be detected (Bertrand et al. 1998).

5.4.1.2 The role of myosin V in synaptic transmission

Nerve cells contain a specialised class of secretory vesicles called synaptic vesicles. These can store small neurotransmitter molecules such as acetylcholine, γ-aminobutiric acid, and others, and mediate the rapid signalling from one cell to another. The regulated neuronal exocytosis involves a cascade of protein-protein interactions. After the synaptic vesicles are filled with neurotransmitter they are transported to the synaptic membrane and docked at defined sites. When an action potential arrives at the nerve terminal it causes an increase in the cytosolic calcium which triggers the vesicles to *fuse* to the membrane and release their contents into the synaptic cleft. The neurotransmitters then diffuse across the synaptic gap and trigger the adjacent nerve cell. The empty vesicles undergo endocytosis and are locally recycled.

A number of proteins are essential for membrane fusion and are conserved in phylogenetically diverse species. SNARE proteins (Soluble NSF-Attachment protein) are believed to be key participants controlling the vesicular docking and fusion. SNARE (Soluble NSF-Attachment protein) on the plasma membrane bind with other SNARE on the synaptic vesicle to form a docking complex. Synaptobrevin or VAMP (vesicle-associated membrane protein) is a SNARE protein

in neurons. It binds to syntaxin and SNAP-25 (Soluble NSF-Attachment protein) on the plasma membrane.

Ohyama et al. (2001) showed that myosin V form a complex with synaptophysin and VAMP in the presence of calmodulin and Ca^{2+} ions. The binding site to VAMP is the myosin globular tail. In the absence of calmodulin only syntaxin-VAMP or myosin-V complex are formed. This findings suggest that VAMP acts as receptor of myosin V on the vesicles and in the presence of Ca^{2+} the complex between myosin V/VAMP binds to syntaxin.

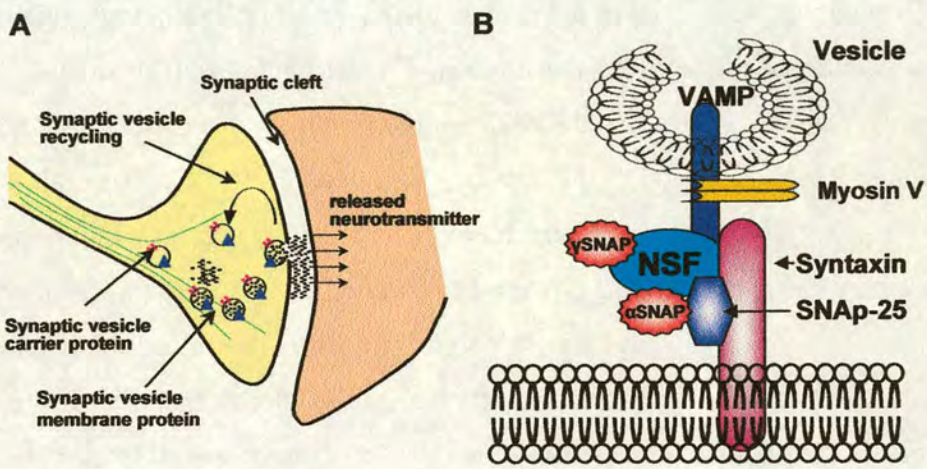


Figure 4.5.1.2 Delivery and docking of synaptic vesicle
A: Formation of synaptic vesicles. The synaptic vesicle is delivered to the nerve synapse where it is loaded with neurotransmitter. The vesicle is then docked to a defined site on the membrane.
B: Proteins needed for correct docking and fusion of the synaptic vesicle to the nerve membrane and the place of myosin V in this complex (see text for details).

5.4.1.3 The role of Myosin V in membrane trafficking

Al-Haddad et al. (2001) showed that Myosin Va is involved in ATP-regulated binding of phagosomes to F-actin. Initially the phagosomes bind to receptors on the cell surface and are enclosed by the plasma membrane. Then they are acidified and fuse with lysosomes containing hydrolytic enzymes which degrade the engulfed microorganisms. These processes together with the subsequent stages of intracellular

transport from the periphery to the centre of the cells are dependent on cytoskeletal proteins. The newly formed phagosomes are surrounded by a thick layer of F-actin and actin binding proteins. The fusion with endocytic organelles is blocked if the cells are treated with drugs that depolymerise F-actin or microtubules. It was found that myosin Va colocalises with fully internalised phagosomes (Swanson et al. 1999). When Al-Haddad et al. (2001) analysed the phagosome movement in normal and dilute macrophages, they observed a rapid accumulation of phagosomes in the perinuclear region.

Using a two-hybrid screen (Lapierre et al. 2001) have identified that rabbit myosin Vb interacts with Rab11a a member of the small GTP-binding proteins in mammals. Additional test showed that myosin Vb tail binds to all members of the Rab11 family: Rab11a, Rab11b, and Rab25 while myosin Va does not bind to any of them. Immunolocalisation and "GFP-myosin Vb" experiments have revealed that myosin Vb is associated with the plasma membrane recycling system in nonpolarised cells and with the apical recycling system in polarised cells. Overexpression of the "GFP-Myosin Vb tail" fusion delayed the transferrin recycling and induced transferrin and transferrin receptor accumulation in pericentrosomal vesicles. These results indicate that myosin Vb is required for transport out of the plasma membrane recycling systems.

5.4.1.4 The role of myosin V in vesicular transport

It has been shown that myosin V (Myo2p) is essential for yeast viability. Myo2p is required for the movement of a portion of the lysosome (vacuole) into the bud and subsequently for its correct inheritance during the division (Catlett and Weisman 1998). Myo2p was first found to deliver the chitosome to the plasma membrane (Santos and Snyder 1997). In addition Myo2p is implicated in post-Golgi vesicle transport and polarised growth.

Class V myosins have a unique tail that contains a coiled coil region and a globular domain which mediate myosin-cargo interactions. Mutations in the globular tail lead to cells with defective vacuole inheritance but with intact polarised growth. A second group of mutations deleting 1459-1491 amino acids of the tail specifically affect polarised growth, thus demonstrating the presence of two distinct regions,

receptors for different cargoes (Catlett et al. 2000). Genetic data suggest that myosin V tail interacts with Smy1p, a kinesin homologue and sec4p, a vesicle associated Rab protein (Schott et al. 1999).

5.4.1.5 The role of myosin V in cell division

The localisation of myosin Va was investigated in cultured primary and immortal mammalian cells (Espreafico et al. 1998; Wu et al. 1998). Immunolocalisation with antibodies specific to myosin Va showed that the protein associate with the MTOC (microtubule organising centre) in interphase cells and to the mitotic asters, spindle and midbody in the dividing cells. In fibroblast cultures prepared from *dilute* cells (null mutant for *Myosin Va)* no such associations were observed. In addition it was found that the fibroblasts in the *dilute* cultures are twice as likely to be binucleated compared to the wild type fibroblasts.

When cells were treated with detergent to remove the membranous structures, myosin Va remained associated with the microtubule-rich domains. Since no Actin was found on these domains it was suggested that myosin V may bind to microtubules either directly or by means of a microtubule associated protein.

5.4.1.6 The role of myosin V in filopodial extension

Myosin V is present in most of the vertebrate tissues but is especially concentrated in nervous tissue and neurosecretory cells. It was shown that myosin V molecules are present in the central and peripheral neurons (Mercer et al. 1991; Espreafico et al. 1992).

To investigate the function of myosin V in growth cone dynamics (Wang et al. 1996) have used chromophore-assisted laser inactivation (CALI) technique. They showed that CALI of myosin-V growth cones of chick dorsal root ganglion neurons increases the rate of filopodial extension while not affecting the rate of filopodial retraction. (Evans et al. 1997) have investigated the distribution of myosin V in superior cervical ganglion (SCG) growth cone. The growth cone contains two well distinguishable cytoplasmic regions: a central microtubule-rich domain and a peripheral domain rich in F-actin. A part of the Myosin V has a diffuse cytoplasmic distribution. It was observed that most of the protein is concentrate in the central

domain. There it appears to associate with small vesicles. In the intermediate region between the central and the peripheral domains myosin V associates with small organelles, actin filament and plasma membranes (Suter et al. 2000). In general organelles associated with myosin V are evident on both microtubules and actin filaments. This suggests that organelles normally transported along microtubules might also be able to move along actin filament using the myosin as a connector molecule. Interestingly no abnormalities were found in growth cones from *dilute* mutants indicating that myosin V is not needed for the traction force for growth cone locomotion or for organisation of the actin cytoskeleton.

5.4.2 Function of Myosin V in *Drosophila*

Myosin V is expressed abundantly during *Drosophila* oogenesis. It subcellular localisation in the ring canal that connects the nurse cells to the developing oocyte implies that this myosin is involved in specific intracellular transport. In addition its accumulation at the anterior pole of the oocyte should be noted. A number of mRNAs and proteins have been shown to localise to the anterior margin of the oocyte. It is possible that Myosin V is involved in the transport and/or docking of molecules or vesicles to the cell membrane.

Myosin V late function during larval development remain to be elucidated. Mutants for *myosin V* demonstrate abnormal morphology and uncoordinated movements. It is likely that Myosin V is involved in vesicular transport of components for the developing CNS (central nervous system). It has been shown that Myosin V in vertebrates is involved in the regulation of neural transmission (Ohyama et al. 2001). The observed *myosin V* mutant phenotype is a likely outcome if the function of myosin V in regulation of neural transmission is disrupted.

5.5 Future work

5.5.1 Roles of *myosin V* transcripts

Our analyses revealed that *myosin V* produces multiple transcripts. Most of the transcripts differ at their 5' and 3'UTRs. Currently we know that transcription is initiated at at least two promoter sites and there are two polyadenylation signals. The difference in some of the transcripts arises from alternative splicing. Additional analysis is needed to elucidate the mechanism controlling this differential expression and determine the tissue and temporal expression of the alternative transcripts. This information is needed to understand the effect of *myosin V* on the cellular processes.

5.5.2 Function of *myosin V* during *Drosophila* development

Having specific antibodies to Myosin V will help us to study the subcellular localisation of the protein. Currently we know that a dynamic expression of m*yosin V* occurs during oogenesis. The expression pattern of Myosin V needs to be studied in embryos and larvae where the mutant phenotype is most severe. It is likely that myosin V is involved in vesicular transport of components for the developing CNS during embryonic development. Co-localisation experiments with antibodies recognising known protein involved in these processes are essential.

The phenotype of myosin V mutations in yeast, mice and humans suggests a role for this myosin in vesicular transport. In addition myosin V appears be important for normal central nervous system function. In all studies myosin V has been shown to localise to vesicular or organelle membranes except one case when it is responsible for the restricted localisation of ASH1 mRNA in yeast.

It is believed that motor-cargo interactions require three components: the motor protein, a cargo binding receptor and accessory components. The question is what are the cargoes and receptors for each motor protein, how it is regulated and which subunit of the motor domain binds the cargo. Studies based on co-immunoprecipitation techniques have indicated that myosin V associates with small GTPases like Rab27a or Rab11a. In addition myosin V was found to bind to SNARE protein to form a docking complex. These proteins have homologues in *Drosophila*. To show whether these are potential targets for myosin V constructs containing the

cargo binding domain (myosin tail) are to be made and tested in two hybrid experiments.

5.5.3 Functional analysis in vivo

Since homozygous mutants for *myosin V* are lethal to study its function will require the production of clones of mutant cells. There are different techniques to make germ-line and somatic clones to look at *myosin V* function in the germ-line and follicle cells.

The mutants can be used with other mutated genes to establish genetic pathways and hierarchies. This will show which genes depend on the correct function of *myosin V* for normal expression.

Further analysis of the function of myosin V can be done by mis-localisation of its expression by using transgenic flies. Expressing Myosin V at a wrong place will allow us to look at any defects in the organisation of the oocyte and egg chamber.

Recombinant constructs carrying sense and anti-sense RNA driven by heatshock promoter have been generated. These need to be transformed in flies. Inducing the promoter can disrupt the function of *myosin V* at specific developmental times and in specific cell types.

5.5.4 GFP fusions

To study the subcellular distribution of *Drosophila* Myosin V, its interactions with other molecules (e.g. actin, calmodulin, dynein and cargo molecules), and any rapid temporal and spatial changes in the distribution of the protein *in vivo* we constructed chimerical GFP fusions. Two different constructs were made: in the first the complete ORF of *myosin V* is fused to a GFP tag, in the second only the tail domain is fused to the GFP. In these constructs GFP can be used as a real time reporter for gene expression reflecting the dynamic events during *Drosophila* development.

Chapter Six: Identification and phylogenetic analysis of *Drosophila melanogaster* myosin genes

6.1 Introduction

The past decade has seen a significant increase in research into myosins. A major effort has been put into finding novel members of this family of actin-based motor proteins. More than 16 classes of myosins have been discovered and characterised, and this number is still rising (Hodge and Cope 2000; Sellers 2000). These myosins are often referred to as unconventional (Mooseker and Cheney 1995). The total number of known myosins is 17 if the conventional two-headed filament forming myosin-II are included in the classification. Myosins have been identified in a wide variety of eukaryotic organisms. Some myosin classes are found in phylogenetically diverse organisms, whereas others, which have arisen later in evolution, have been found in a single organism only (Table 6.1).

Current research concentrates on the functional analysis of these new types of myosins. A number of studies suggest that these motors play important roles in a variety of cellular functions including organelle, RNA and protein transport, maintenance of cell architecture, cell movements, and signal transduction.

Table 6.1 Species diversity of the myosin genes

Myosin I	Myosin II	Myosin III
Acanthamoeba castellanii *Saccharomyces cerevisiae* *Dictyostelium discoideum* *Caenorhabditis elegans* *Drosophila melanogaster* *Rana catesbeiana* *Gallus gallus* *Rattus norvegicus* *Bos taurus* *Homo sapiens*	*Acanthamoeba castellanii* *Saccharomyces cerevisiae* *Aequipecten irradians* *Dictyostelium discoideum* *Caenorhabditis elegans* *Drosophila melanogaster* *Xenopus laevis* *Gallus gallus* *Rattus norvegicus* *Homo sapiens*	*Limulus polyphemus* *Drosophila melanogaster* *Homo sapiens*
Myosin IV	**Myosin V**	**Myosin VI**
Acanthamoeba castellanii	*Saccharomyces cerevisiae* *Caenorhabditis elegans* *Drosophila melanogaster* *Gallus gallus* *Mus musculus* *Rattus norvegicus* *Homo sapiens*	*Caenorhabditis elegans* *Morone saxatilis* *Drosophila melanogaster* *Mus musculus* *Sus scrofa* *Homo sapiens*
Myosin VII	**Myosin VIII**	**Myosin IX**
Dictyostelium discoideum *Caenorhabditis elegans* *Drosophila melanogaster* *Mus musculus* *Sus scrofa* *Homo sapiens*	*Arabidopsis thaliana* *Helianthus annuus* *Zea mays*	*Caenorhabditis elegans* *Mus musculus* *Rattus norvegicus* *Homo sapiens*
Myosin X	**Myosin XI**	**Myosin XII**
Mus musculus *Bos taurus* *Homo sapiens*	*Chlamydomonas reinhardtii* *Chara corallina* *Arabidopsis thaliana* *Helianthus annuus* *Zea mays*	*Caenorhabditis elegans*
Myosin XIII	**Myosin XIV**	**Myosin XV**
Acetabularia cliftonii	*Toxoplasma gondii* *Plasmodium falciparum*	*Drosophila melanogaster* *Mus musculus* *Homo sapiens*
Myosin XVI	**Myosin XVII**	**Myosin XVIII**
Rattus norvegicus	*Aspergillus nidulans* *Magnaporthe grisea*	*Drosophila melanogaster* *Mus musculus* *Homo sapiens*

6.1.1 Myosin domains

6.1.1.1 Myosin motor domain

All known myosins comprise an N-terminal head domain, a neck regulatory domain, and a specific carboxy-terminal tail domain (Figure 6.1.1.2) (Mooseker and Cheney 1995). The head or motor domain contains ATP and actin binding sites, and is responsible for the mechanochemical properties of the protein (Gilbert and Mackey 2000). In addition the head contains three loops and two switches (see appendix I). The first loop is responsible for the speed at which myosin moves, loop 2 is part of the actin binding interface and is believed to control the rates of phosphate release and to set the maximum speed for ATPase activity. Switch 1 and switch 2 have been shown to control the rate of ATP hydrolysis (Onishi 1998). Myosins show an actin-stimulated Mg^{2+} ATPase activity, thus converting the energy stored in ATP into mechanical force (Volkmann and Hanein 2000). The latter is used to move the myosin molecules along actin filaments or translocate other molecules (Hasson and Mooseker 1995; Langford 1995).

6.1.1.2 Myosin neck domain

- ***IQ motifs***

The neck domain contains regulatory sites, composed of IQ (isoleucine /glutamine) motifs, repeats of 23-30 amino acids (Mercer et al. 1991; Rhoads and Friedberg 1997; Friedberg and Rhoads 2001). The consensus sequence for the IQ motif is **IQ**xxx**RG**xxx**RR** ("x" is any residue). Each IQ motif binds to light chains that are related to the EF-hand family of proteins that include calmodulin (CaM) and troponin C (Kawasaki et al. 1998). The size of the neck domain varies from one to seven IQ tandem repeats (Figure X) and can vary among members within a particular class. In addition the neck is often the site of alternative splicing. This produces necks with variable length (variable number of IQ repeats), which are associated with regulatory function. The number of the IQ motifs is unrelated to myosin classification.

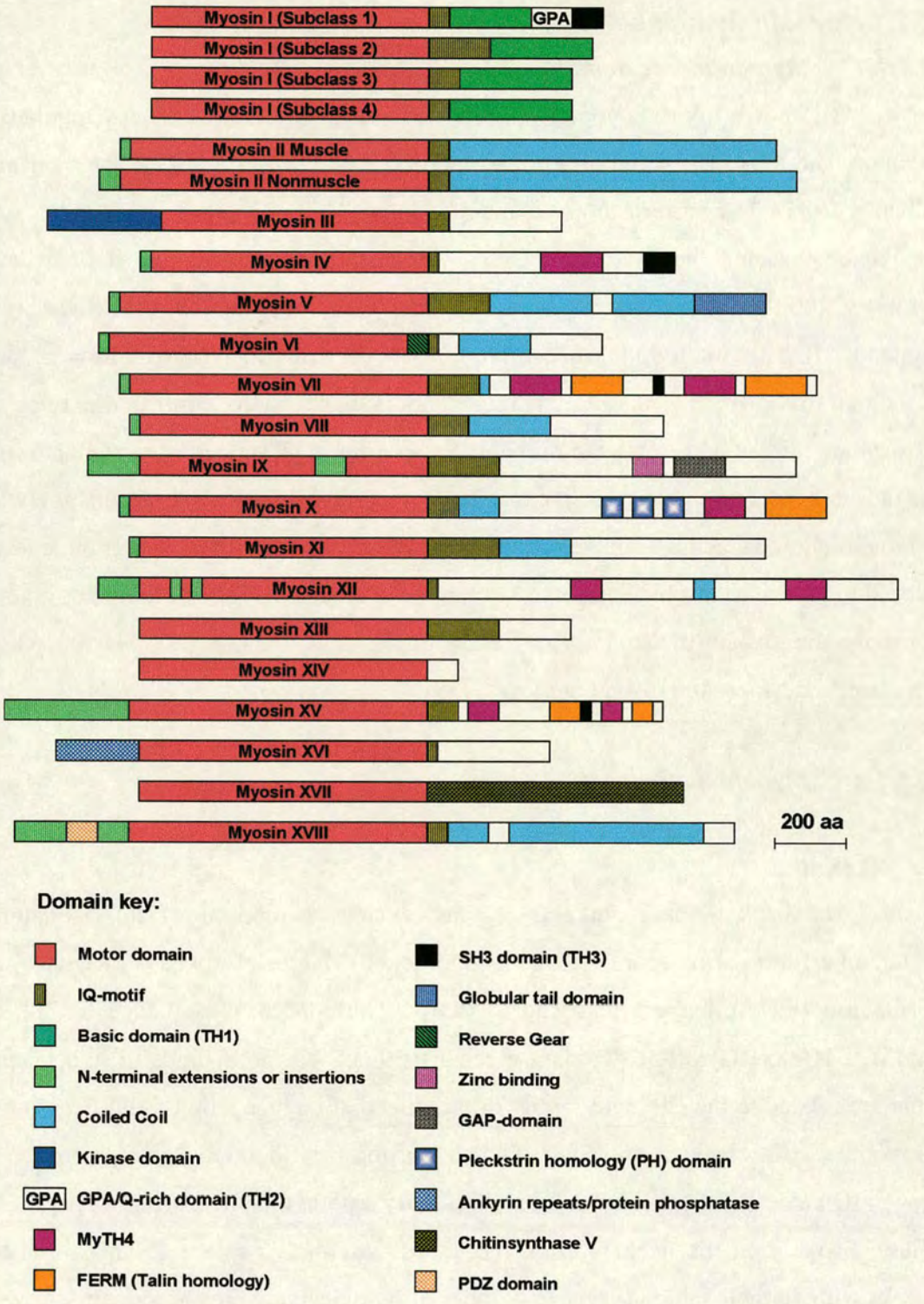


Figure 6.1.1.2 Domain structure of conventional and unconventional myosins

The domains are colour coded and explained in the figure legend. For more information see the text .

It has been proposed that the neck domain associated with its light chains serves as a lever arm for force transduction and determines the step size with which the myosin molecule moves forward along the actin filaments (Highsmith 1999).

- ***Myosin light chains***

Myosins from class II have two IQ motifs in their neck domain and bind two light chains. The First IQ motif (proximal to the motor domain) binds an essential light chain (ECL) and the second IQ motif binds regulatory light chain (RLC). Most of the unconventional myosins associate with one or more copies of CaM as light chains. Unconventional myosins with up to six IQ motifs have been shown to bind 4-5 CaM light chains and 1-2 class specific light chains (Espindola et al. 2000; (Stevens and Davis 1998).

Drosophila has two genes encoding ELC, one muscle specific and one nonmuscle specific light chains (Edwards et al. 1995; Edwards and Kiehart 1996). The nonmuscle isoform associates with nonmuscle myosin II and is thought to bind to myosin V, by analogy with the chicken myosin V (Espindola et al. 2000). In addition there two RLC: a non muscle isoform encoded by *spaghetti-squash, (sqh)* gene and a muscle isoform expressed in all muscle tissues (Kiehart and Feghali 1986). The activity of RLC is regulated by phosphorylation (Karess et al. 1991). The phosphorylation site is usually a serine residue located near the N-terminus which often is accompanied by a second (one amino-acid apart) threonine phosphorylation site. Only one calmodulin gene has been found in *Drosophila.* There is also a testes specific protein, androcam, that shows a limited similarity to calmodulin. Analysis of the recently annotated *Drosophila* genome identified at least one other calmodulin-like gene (Yamashita et al. 2000).

- ***Calmodulin signalling via the IQ motif***

In general, the light chains activates a diverse group of target cellular proteins when bound to Ca^{2+}. These IQ-containing proteins include spindle associated proteins, neuronal growth proteins, conventional and unconventional myosins, ion channels, phosphatases, Ras and Ras-GAP like proteins. The majority of IQ-containing proteins bind calmodulin in the presence of Ca^{2+} (calcium-dependent

regulation). In this way the Ca^{2+} signal is transferred to the calmodulin to control the molecular motor activity, cytoskeletal organisation, mitosis, etc. (Bahler and Rhoads 2002).

Interestingly most of the unconventional myosins carry IQ-motifs that bind calmodulin with higher affinity in the absence of Ca^{2+}. This is due to conformational changes of the calmodulin upon Ca^{2+} binding but not due to dissociation of calmodulin from the myosin. These conformational changes modulate the myosin ATPase activity and are not just on/off switches (Bahler and Rhoads 2002; Homma et al. 2000).

6.1.1.3 Myosin tail

Following the neck domain, each myosin has a highly divergent tail domain. A subset of myosin tails have predicted coil-coil α-helical domains, which promote dimer formation; a typical example being the two-headed conventional myosin-II. Some other myosins lack coiled-coil domains but contain structural domains found in other proteins (Figure 6.1.1.2 and Table 6.1.1.3).

6.1.2 Myosin classification

The classification of myosins is based on the sequence comparison of their core motor domains (myosin head), equivalent to amino acids 88 to 780 of chicken skeletal myosin II (Cope et al. 1996). The motor domain is highly conserved among all myosins, reflecting the high conservation of its function. However, they have a number of class specific features (characteristic inserts or substitutions) which might be important in defining the precise function of a given myosin (Table 6.13). Phylogenetic analysis of the tail domain sequences produces similar results, indicating that heads and tails have co-evolved (Korn 2000).

6.1.3 Structure and function of conventional and unconventional Myosins

All myosins are composed of a head domain, neck and a class specific tail domain. It is the tail domain that is largely responsible for the class-specific functions (Table 6.1.3A). It has been shown that myosins are involved in a variety of cell functions such as organelle and mRNA transport, endo- and exocytosis, cell

division, maintenance of the cell architecture, regulation of the actin assembly and a role in the dynamics of actin-rich cell extensions, coordination and cooperation between microtubule- and actin dependent transport and signal transduction. Recent studies have identified a variety of segments (motifs) in the myosin tail domains (for review see (Oliver et al. 1999). Some of these are predicted to form coiled-coil domains that promote homodimerisation of two myosin heavy chains. Some remain with unknown function. Others showed high homology to domains with known function found in other proteins (Table 6.1.3B).

6.1.3.1 Myosin I

Class I myosins have a variable number of IQ domains (one to six) in their neck domain. They lack coiled-coil domains in the tail region. Instead, following the tail domain there are three tail homology (TH) domains. TH1 has been shown to bind to phospholipid membranes and implicated in targeting myosin chains to membranes. TH2 domain or GPA/GPG is a proline-rich domain. It contains an actin-binding motif that is ATP insensitive. The TH3 domain is believed to promote binding to proline rich proteins (Osherov and May 2000; Barylko et al. 2000). Studies have shown that actually it is the SH3 motif found in the TH3 region that is important for actin binding. It was shown that the SH3 domain is able to induce cytosol-dependent actin-polymerisation in yeast (Geli et al. 2000).

Ameboid class I myosins have role in actin-dependent movements and are responsible for the shape of a variety of subcellular structures such as vacuoles, endosomes and cell protrusions like pseudopodia (Ostap and Pollard 1996; Vargas et al. 1997; Senda et al. 2001).

In vertebrates myosin Iβ was found to localise primarily to the tip of stereocilia protrusions that are found on the top of hair cells (Hasson et al. 1997; Barylko et al. 1992) Overexpression of myr1 in NRK (Normal rat kidney) cells showed no phenotype suggesting its function is not essential. In contrast overexpression of BBMI (brush border myosin I) in hepatoma cells caused defects in endocytosis and structural changes, while overexpression in epithelial cells disrupted the traffic of basolaterally internalised molecules to the apical surface (Osherov and May 2000).

Table 6.1.3A Domain structure and function of myosin classes

Class Myosin	IQ	Hc*	Other (N- and C terminal) Domains	Function
I Subclass 1	1 -2	1	TH1, GPA/GPQ and SH3	Vesicle transport, cell growth and cell motility.
Subclass 2	3-6	1	TH1 domain	Function in the microvilli of the brush border.
Subclass 3	3	1	TH1 domain	Epithelial morphogenesis and hair cells function.
Subclass 4	2	1	TH1 domain	Epithelial morphogenesis.
II Muscle	2	2	CC	Smooth or skeletal muscle contraction.
Nonmuscle	2	2	CC	Maintenance of the cell architecture, cell motility and phagocytosis.
III	1-2	1	N-terminal protein kinase	Role in prototransduction.
IV	1	1	MyTH4 and SH3	? (Only in *Acanthamoeba* species).
V	6	2	CC, Transmembrane, specific, DIL domain	Membrane trafficking, polarised cell growth. Vesicle, protein and/or mRNA transport.
VI	1	2 ?	CC, Reverse gear, specific, Globular tail domain	Vesicle transport, epithelial morphoge-nesis, and stereocilia function. Spermatid individualisation. Moves toward the "-"end of actin filaments.
VII	4-5	2 ?	CC, MyTH4, FERM, and SH3	Membrane trafficking, hair and photoreceptor cells function.
VIII	3-4	2 ?	CC, Serine-rich domain, C-terminal domain	Cell wall function in plants, intracellular transport.
IX	4-6	1	N-terminal extension, Zinc binding and Rho-GAP	Signalling (GTPase activating).
X	3	2 ?	CC, PH, MyTH4, FERM	Localised to regions of dynamic actin. Signal transduction.
XI	5-6	2 ?	CC	Vesicular transport in plants.
XII	IQ ?	2 ?	CC, N-terminal, MyTH4	? (In *C. elegans* only).
XIII	4-7	1	-	? (In plants only).
XIV	IQ ?	1	-	? (In *Toxoplasma* and *Plasmodium* species)
XV	2-3	1	N-terminal extension MyTH4, FERM, and SH3	Hair cell function
XVI	2	1	Ankyrin repeats	Neuronal cell migration
XVII	IQ ?	1	Chitin synthase domain	? (In *Pyricularia* and *Emiricella* sp.)
XVIII	1-2	2 ?	CC, KE, and PDZ domain	? Maintenance of the stromal cell architecture

Note: *The number of heavy chains reflects their ability to dimerise (based on coiled-coil predictions). Key to domain abbreviations: TH1- Tail Homology basic 1 domain, GPA-glycine/proline/alanine rich domain, GPQ-glycine/proline/glutamine rich domain, SH3-Tail Homology 3 domain (binds to proline-rich motifs), CC-Coiled-Coil, MyTH4-Myosin Tail Homology 4 domain, Rho-GAP-domain (activates small GTPases of the Rho family), PH-Pleckstrin Homology domain, KE-lysine/glutamate rich domain, PDZ or DHR (Dlg homologous region . For recent reviews on the structure and properties of the unconventional myosins see (Wu et al. 2000; Oliver et al. 1999; Baker and Titus 1998; Mermall et al. 1998). In the construction of this table data was used from these reviews, the myosin home page and Cope et al. (1996).

Table 6.1.3B Function of myosin domains

Domain	Full name	Function	Distribution
MYSc	Myosin motor	Binds actin and ATP. Hydrolysis ATP to generate mechanical force.	All myosins
IQ motif	-	Binds calmodulin and/or other light chains of the EF-hand family.	All myosins, except XIV, XVII
CC	Coiled-coil	Responsible for dimerisation. Involved in the assembly of myosin thick filaments.	In myosins II, V,VI, VII, VIII, X, XI, XII, XVIII
TH1	Tail homology basic 1,	Involved in membrane binding.	Myosin I
GPA (TH2)	Tail homology 2 (Gly-Pro-Ala-rich)	Role in actin binding in an ATP-insensitive manner.	Myosin I
GPQ (TH2)	Tail homology 2 (Gly-Pro-Glu-rich)	Role in actin binding in an ATP-insensitive manner.	Myosin I
SH3 (TH3)	Tail homology 3	Involved in signal transduction. Binds proline-rich motifs.	Myosin I, IV, VII and XV
MyTH4	Myosin tail homology 4 domain	Unknown	Myosin IV, VII, X, XII and XV
Kinase domain	-	Binds ATP. Signal transduction from the photo-complex to the cytoskeleton.	Myosin III
DIL	Globular tail with AF-6 homology	Function unclear. Implicated in cargo binding.	Myosin V
FERM	Band 4.1, ezrin, radixin, moesin homology	Possible role in connecting the cytoskeleton to the plasma membrane.	Myosin VII, X and XV
Reverse gear	-	Directs a movement towards the minus end of actin filaments	Myosin VI
DAG-PE-bind	Zinc binding	Binds Zn^{2+} ions. Though to bind to diacylglycerol, a protein kinase C activator.	Myosin IX
Rho-GAP	-	Activates small GTPases of the Rho family (signal transduction). Reorganisation of the actin cytoskeleton.	Myosin IX
Rho GEF	-	Activates small GTPases of the Rho family (signal transduction). Known to associate with PH domains.	*Dictyostelium* MyoM
RA	(RAS Association)	Known to associate with RasGTP. Unknown function in the myosins.	Myosin IX
PH	Pleckstrin homology	Known to bind to: G proteins, lipids, phosphatidylinositol, phosphorylated Ser/Thr residues and membranes.	Myosin X
Ank	Ankyrin repeats/protein phosphatase	Adaptors, associating with spectrin-based cytoskeleton and membrane proteins. Protein/protein interactions.	Myosin XVI
Chitin synthase	-	A role in cell wall synthesis.	Myosin XVII
PDZ	Dlg homologous region	Signal transduction. Targeting signalling molecules to sub-membranous sites.	Myosin XVIII
KE	Lysine-glutamate-rich	Unknown function	Myosin XVIII
WD40	Similar to motifs found in beta-transducing subunits of G proteins	Unknown function	Plasmodium Myosin

6.1.3.2 Myosin II

Myosins from class II were the first members of the myosin superfamily to be identified. They have a long α-helical domain in the tail that allows formation of homodimers. The dimers can further assemble into bipolar anti-parallel filaments.

Lower eukaryotes like *Dictyostelium* and *Acanthamoeba* species have relatively divergent types of myosin II. They have been implicated in a variety of functions such as receptor capturing, generation of cortical tension and cytokinesis (Cheney et al. 1993b).

In vertebrates myosin-IIs are divided into two major groups: striated the muscle myosin and the nonmuscle (and smooth muscle) group. Skeletal muscle myosin is responsible for generating the force for muscle contraction. In skeletal muscle myosin II filaments are surrounded by thin actin filaments in complex arrays. Thus ATP-driven sliding of the actin filaments results in muscle contraction. Nonmuscle myosins have role in cytokinesis and capping surface receptors (Boyne et al. 2000), cytoplasmic transport, maintenance of the cell architecture, cell motility and phagocytosis (Edwards and Kiehart 1996; Ikonen et al. 1997).

6.1.3.3 Myosin III

A member of this class was first identified from *Drosophila* and named *ninaC* (Montell and Rubin 1988). NinaC myosins have also been found in human retina (MYO3 and MYO3B), fish retina (FM3a and FM3b) (Berg et al. 2001) and (Dose and Burnside 2002) and in horseshoe crab, *Limulus polyphemus* (Battelle et al. 1998).

Nina C protein has ~300 amino-acid kinase domain located on its N-terminus. The gene has been shown to express two alternatively spliced forms p132 and p174 the larger of which contains a rhabdomere localisation signal (Montell and Rubin 1988; Hicks and Williams 1992). NinaC expression is detected only in photoreceptor cells. The protein localises to the tip of the rhabdomere, that is a site of phototransduction. There ninaC binds to inaD scaffolding protein and links it to the actin filament at the core of the rhabdomere. InaD is a key scaffolding protein which associates via its five PDZ domains with most of the proteins responsible for the phototransduction signalling (rhodopsin, phospholipase C, protein kinase C and TRP

and TRPL cation channels) to form a signalling complex known as "transducisome" (Wes et al. 1999). Defects in *ninaC* leads to abnormal electrophisiological response and retinal degradation, indicating a need for phosphorylation in regulating the termination of photoresponse.

6.1.3.4 Myosin IV

Currently M4 from Acanthamoeba is the only member of this class (Horowitz and Hammer 1990). M4 does not have coiled-coil motifs and does not form dimers. The protein contains a single IQ motif thought to bind calmodulin. The tail domain contains one SH3 domain which can bind to proline-rich motifs. A proline-rich ligand, Acan 125 has been identified for the SH3. Acan 125 belongs to the superfamily of leucine-rich repeat proteins (Xu et al. 1997). In addition the tail has one MyTH4 domain with unknown function (Oliver et al. 1999).

6.1.3.5 Myosin V

This type of myosin is extensively discussed in chapter 3.

6.1.3.6 Myosin VI

Drosophila myosin 95F was the first member of this family to be described (Kellerman and Miller 1992). Later two genes were identified in *C. elegans* (Baker and Titus 1997) and a single gene in vertebrates (pig, mouse, chicken and human). Analysis of the myosin VI structure shows that the protein has a 25 amino-acid insertion in the conserved head domain and a threonine substitution at the TEDS rule site. This suggests that myosin function is regulated by head phosphorylation. The neck domain contains a single IQ motif preceded by a unique 53 amino-acid reverse gear (RG) domain. The latter has a role in reversal of lever arm movements. Analysis of the sequence showed that RG domain is positioned after the converter region, at the base of the long α-helix that forms the backbone of the lever arm. This shows that the insert (RG) does not replace the converter but modulates its function to redirect lever arm movement (Hasson and Cheney 2001). The myosin VI tail contains a coiled-coil domain, responsible for dimerisation and a distal globular domain. Recent studies have shown that the globular tail associates with microtubule-binding protein CLIP-170 (Lantz and Miller 1998) demonstrating the presence of dynamic

interaction between components of the microtubule- and actin- based transport systems. Deng et al. (1999) have shown that Myosin VI is required for follicle cell epithelial morphogenesis during *Drosophila* oogenesis. It was shown that *Drosophila* Myosin VI has a role in sperm individualisation (Hicks et al. 1999).

Class VI myosins are one of three classes (the others are class VII and XV) that are associated with deafness. Mutations in mouse myosin VI (Snell's waltzer mouse) result in deafness and vestibular problems (Avraham et al. 1995). Both defects are associated with abnormal development of the organ of Corti. The latter projects more than a 100 actin-packed stereocilia from its apical surface.. The base of each stereocilia is anchored into the cuticular plate. When the stereocilia bundles are deflected due to interaction with sound waves, this result in stretching the links between the stereocilia in different rows which opens transductions channels. This in turn leads to depolarisation of the hair cell and release of neurotransmitter (Holme and Steel 1999). The Snell's waltzer mice lack hair cell stereocilia and are unable to engage in mechanotransduction.

6.1.3.7 Myosin VII

Myosin VII has a typical structure for the unconventional myosins. The molecule has a neck domain composed of five IQ motifs and a very long tail domain. The tail contains a number of motifs that are believed to interact with other proteins: a short coiled-coil domain, which has been shown to promote homodimerisation, two MyTH4 domains, two FERM domains and a poorly conserved SH3 domain. The FERM domains of these proteins have been shown to bind membrane proteins (directly or via an adapter).

Mutations in myosin VIIa cause the *shaker-1* phenotype in mice and Usher syndrome 1B in humans (Friedman et al. 1999). These phenotypes are characterised by deafness and vestibular dysfunction. In addition the mutants exhibit abnormal microtubule organisation in the sperm-tail axonemes, photoreceptor cells, and nasal cilial cells. The expression pattern of myosin VII in the sensory hair cells in both the cochlea and vestibular apparatus correlates well with the sites that are affected in the pathogenesis (Oliver et al. 1999; Holme and Steel 1999). In addition myosin VII is expressed in a variety of organs: brain, intestine, liver, kidney, adrenal gland and

testis. Interestingly an abnormal phenotype was only observed in the inner ear and eye.

Recent studies have demonstrated a role for myosin VII in phagocytosis in *Dictyostelium* amoeba (Titus 1999).

6.1.3.8 Myosin VIII

Members of this class are plant myosins. Myosin class VIII have a large coiled-coil motif in their tail known to promote dimer formation. Immunolocalisation experiments have shown that *A. thaliana* myosin 1 (class VIII) appears to be localised in the plasmodesmata in the newly formed cell walls. It was suggested that myosin VIII is involved in the maturation of the cell plate and restablishment of the actin network at sites of intercellular communication (Reichelt et al. 1999).

6.1.3.9 Myosin IX

Myosins from class IX have been isolated from a number of species. Analysis of the genomic sequence of *C. elegans* indicates that it has a single gene for myosin IX (Baker and Titus 1997). Vertebrates were shown to have two types of myosin IX: a and b subtypes. Myosins *MYO9A* and *MYO9B* have been cloned from human (Wirth et al. 1996; Berg et al. 2001). In rat M9b was cloned first and is known as *myr 5*. Recently the second member, M9a has been identified and named *myr 7*. (Reinhard et al. 1995).

The heads of these myosins have a ~60 amino-acid extension and large insertions of approximately 120 amino-acids in the actin-binding region. The tail lacks any coiled-coil motifs suggesting that the myosin functions as a monomer. The tail contains regions found in the chimerins, with a putative C_6H_2 zinc binding domain and a GAP (GTPase activating protein) domain of the rho/rac family of ras-like G-proteins. The ras-related rho subfamily of small GTP binding proteins (Rho, Rac and Cdc42Hs) has a role in actomyosin-dependent processes and in actin reorganisation. It was demonstrated that *MYO9b* and *myr 5* exhibit GAP activity for rho but almost no activity for rac or Cdc42Hs. Transgene expression of *myr 5* in insect cells resulted in variety of changes in the thin processes (numerous processes of rapid growth and retraction during dendritic development). Overexpression

experiments in mammalian cells caused cell rounding and loss of actin stress fibers (Oliver et al. 1999; Chieregatti et al. 1998) have shown that *myr 5* is expressed mainly in lung, spleen, liver and testis, while *myr 7* (MYO9a) is abundantly expressed in brain. It is believed that it plays a role as a negative regulator of rho-mediated actin morphology and function in neurons.

6.1.3.10 Myosin X

Myosin X was first isolated from a bovine aorta library. Later members have been identified in human and mouse (Hasson et al. 1996; Oliver et al. 1999). Myosins class-X have a neck with three IQ motifs, short coiled-coil domain, three plextrin homology domains (PH) followed by MyTH4 and FERM domains. All these domain are known to have role in protein-protein interactions. The presence of all these domain in a single molecule makes it a potential candidate to form a part of a multicomponent signalling complex.

6.1.3.11 Myosin XI

This class contains a number of plant genes which were previously classified as class-V myosins due to structural similarities with the animal class-V myosins (Liu et al. 2001). As such these were discussed in chapter 3. It was shown that that plant myosin XI localises to the root tip cells and is highly concentrated in mitochondria and low-density membranes. It is believed that myosin XI has a role in vesicular transport and membrane trafficking.

6.1.3.12 Myosin XII

This class is defined by the *hum-4* gene isolated from *C. elegans* (Baker and Titus 1997). The gene encodes an unusually large heavy chain of 307kDa. Phylogenetic analysis showed that the head domain show only 20% identity to other myosins. The function of HUM-4 remain unknown. Sequence analysis has shown the presence of two inserts in the head domain, two MyTH4 domains and a small coiled-coil domain implicated in dimer formation.

6.1.3.13 Myosin XIII

There is little information about these myosins. So far only two genes have been isolated, both from the green algae *Acetabularia cliftonii*. Myosin XIII has a large neck domain composed of up to seven IQ motifs implying a complex regulation. The tail domain is short and contains no domain with known function.

6.1.3.14 Myosin XIV

These myosins are found only in obligate intracellular parasites such as *Plasmodium falciparum* and *Toxoplasma gondii* (Hettmann et al. 2000). The P*lasmodium* expresses a multiple of short tailed myosin XIV genes: TgMA, TgMB and TgMC, while only a single gene have been identified in *T. gondii so* far. It has been shown that class-XIV myosins are targeted to the cell periphery where they associate with the membrane. This localisation is actin independent. Site directed mutagenesis revealed that two arginine residues in the distal part of the tail are responsible for the specific membrane localisation. A function in the invasion process has been suggested for these proteins.

6.1.3.15 Myosin XV

Myosins from this class have been identified in mouse, human and recently in *Drosophila* (Liang et al. 1999; Tzolovsky et al. 2002). The human (*MYO15*) and mouse (*Myo15*) encode approximately a 360 kDa heavy chain that is unique among the myosins. The protein has two to three IQ motifs followed by a very long tail domain composed of two MyTH4 domains, two FERM domains and a single SH3 domain. Myosin XV is abundantly expressed in the pituitary gland and the sensory epithelia of the cochlea and vestibular system in the inner ear (Lloyd et al. 2001; Friedman et al. 1999). Mutations in myosin XV are responsible for congenital profound deafness *DFNB3* in humans and deafness and abnormal behaviour due to defects in the vestibular system in shaker-2 mice.

6.1.3.16 Myosin XVI

A single member of this class has been isolated from rat and designated as *myr 8* (Patel et al. 2001). The gene expresses two alternatively spliced isoforms, 8a and 8b. The head of myosin XVI contains a large N-terminal extension that contains

multiple ankyrin repeats. They are known to associate with protein phosphatase 1 (PP1) catalytic subunits 1α and 1γ. The neck containing a single IQ motif is followed by a short tail that is positively charged in 8a and bearing no charge in 8b. It was shown that 8b is the predominant isoform expressed in brain. It is believed that myosin XVI selectively targets protein phosphatase catalytic subunits to the developing brain.

6.1.3.17 Myosin XVII

The first member of this class was cloned from *Aspergillus nidulans* (Fujiwara et al. 1997). The gene was named *csmA* for chitin synthase with a myosin motor-like domain. It encodes a large chitin synthase domain situated at the C-terminus. No putative IQ motifs have been identified so far in the neck domain. Recently a second member has been found in *Pyricularia oryzae* (Park et al. 1999). It was suggested that myosin XVII has a role in directing the chitin synthesis.

6.1.3.18 Myosin XVIII

Using a differential display method Furusawa et al. (2000) have isolated a gene with expression levels correlated with the hematopoetic supportive ability of the stromal cells. The gene proved to be a novel member of the myosin family. Recent studies have identified myosin XVIII genes in human (Berg et al. 2001) and *Drosophila* (Tzolovsky et al. 2002). Myosin XVIII contains a lysine/glutamine-rich sequence (KE sequence) followed by a single PDZ domain. The neck contains one or two IQ domains. The tail has a long coiled-coil domain similar to that of class-II myosins, followed by a unique, 100 amino-acid globular tail. Currently the function of myosin XVIII remain unknown. It is believed that it functions as a part of a mobile scaffolding complex through interactions with the PDZ domain. In general the role of PDZ domains is to position correctly target ion channels, receptors or other signalling molecules to specialised sites in the cell (Bezprozvanny and Maximov 2001).

6.1.4 Myosin genes in eukaryotic organisms

Five myosin genes have been identified in yeast (*S. cerevisiae*) falling into three classes: there are two class I myosins, one class II, and two class V myosins (Brown 1997). It was suggested that the whole yeast genome has undergone a duplication in ancient times followed by a number of modifications. As a result a small fraction of the genes were retained in duplicate (most being deleted), thus explaining the loss of the second myosin II gene (Wolfe and Shields 1997). *S. cerevisiae* is the organism with the lowest known number of myosin genes. This demonstrates that a eukaryote can function with a set of only three types of myosins.

Multicellular organisms have the ability to express some 10 to 40 myosin genes encoding at least six types of myosin. It seems that multicellular organisms require many more specialised myosins than the single cell yeast. This has raised the question of what is the degree of functional redundancy between the classes and between the members of a given myosin class?

So far 11 myosin genes have been identified in the slime mould *D. discoideum*. Despite the fact that this is one of the simplest multicellular organisms, it expresses a diverse set of myosin genes (Soldati et al. 1999). They encode at least six different classes of highly specific myosins. There are six class I myosins (MyoA, B, C, D, E, K and probably MyoF), a single member of class II (MhcA), class VII (MyoI) (Titus 1999) and class XI (MyoJ) (Hammer and Jung 1996). The highly divergent MyoM is still to be classified (Schwarz et al. 1999).

Fourteen myosin genes have been identified in the nematode *C. elegans* (Baker and Titus 1997). They encode two structurally distinct class I myosins, six class II, one class V, two class VI, one class VII and one class IX myosin. It was found that *C. elegans* has a highly divergent type of myosin, being the founding and only member of class XII myosins.

The situation with vertebrates appears even more complex. They express some 40 myosin genes grouped in 12 classes. In human there are eight class I, sixteen class II, two class III, three class V, one class VI, two class VII, two class IX, one class X, two class XV, and one class XVI myosins (Hasson et al. 1996; Berg et al. 2000; Berg et al. 2001). Recent studies have discovered two PDZ-containing

myosins (Furusawa et al. 2000) as well as a novel unclassified myosin (Berg et al. 2001).

6.1.5 Myosin genes in Drosophila

In *Drosophila* eight different myosin genes were described thus far (Fig. 6.1.5, Fig. 6.2 and Table 6.3). There are two class I myosins, members of subclass 3 (myosin-IB) and subclass 4 (myosin-IA) (Morgan et al. 1994; Cheney et al. 1993a; Mooseker and Cheney 1995). Only a single muscle myosin II gene was found in *Drosophila* (Hastings and Emerson 1991; Bernstein and Milligan 1997). It encodes more than 13 protein isoforms with complex temporal and spatial expression patterns (Bernstein and Milligan 1997; Zhang and Bernstein 2001). *Drosophila* has a second myosin-II gene that encodes a cytoplasmic nonmuscle myosin (Kiehart et al. 2000; Mansfield et al. 1996). The founding member of class III myosins was discovered in *Drosophila* (Montell and Rubin 1988). The *ninaC-III* gene encodes two isoforms resulting from alternative RNA splicing. These differ in the composition of their C-terminal tails and show differential expression patterns (Porter et al. 1992; Li et al. 1998). A single myosin V gene, with at least two different splice forms, was identified in *Drosophila* (Bonafe and Sellers 1998; MacIver et al. 1998). Kellerman and Miller (1992) cloned a novel unconventional myosin from *Drosophila* - the first member of class VI myosins. The gene encodes multiple protein isoforms, which are present throughout *Drosophila* development (Mermall and Miller 1995; Deng et al. 1999 Hicks et al. 1999). *Drosophila* myosin VIIA was the first member of this class to be described (Cheney et al. 1993b; Kiehart et al. 1998).

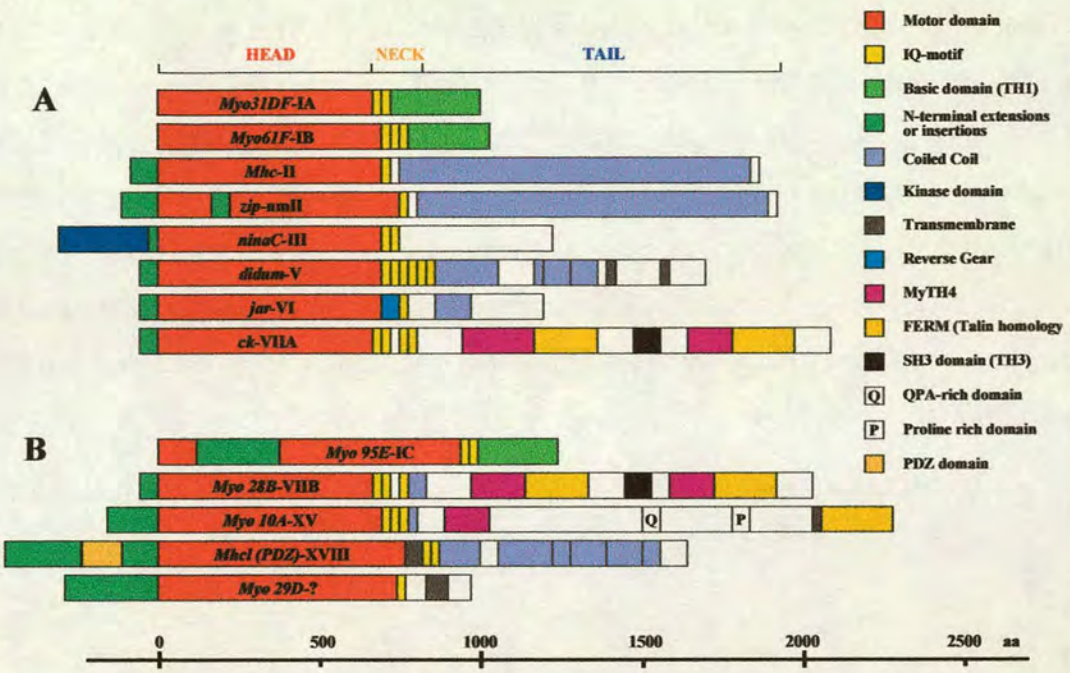


Figure 6.1.5 Schematic diagram of the domain structure of *Drosophila* myosins

A: The previously described *Drosophila* myosins. Their sequence and structure have been studied and confirmed. For DNA and protein accession numbers see Table 6.3

B: The five new or previously uncharacterised *Drosophila* myosins. Their domain structure is based on the predicted cDNAs for these myosins (see Table 6.3). The presence of some domains (e.g. the PDZ domain in Mhcl myosin) was verified with RT-PCR. A colour key to the domain names is given on the left. Some of the names are abbreviated, these being: MyTH4-Myosin Tail Homology 4 domain, FERM (Band 4.1, ezrin, radixin, moesin-homology domain), QPA-glutamine/ proline/alanine rich domain, PDZ- a domain with potential for targeting to membrane sites.

6.2 *Analysis and Manipulation of Sequences*

We have used the completed *Drosophila melanogaster* Genome Project to determine the number of myosin encoding genes in this species and to classify them. The new myosin genes were identified by comparing the *Drosophila* genome sequence with the conserved head (equivalent to amino acids 88 to 780) of chicken skeletal myosin II. Comparison to the Genbank Data Base was done using the BLASTP algorithm on the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/BLAST/), the Pôle Bio-Informatique Lyonnaism

server (http://pbil.univ-lyon1.fr/BLAST/blast.html), and The Berkeley Drosophila Genome Project BLAST at http://www.fruitfly.org/blast/. This search retrieved the eight previously described myosin genes as well as three new genes: *Myo 28B1* (AAF52536), *Myo10A* (AAF47980), and *Mhcl* (or *Myo89B*) (AAF55271). Subsequent detailed analysis was done using the conserved head or specific tails from all known classes of myosins to search the translated *Drosophila* genome sequence (TBLASTN on the NCBI server). These searches identified two other sequences highly similar to myosins: *Myo29D* (AAF52683, AF454348) and *Myo95E* (AAF56246, AF454350).

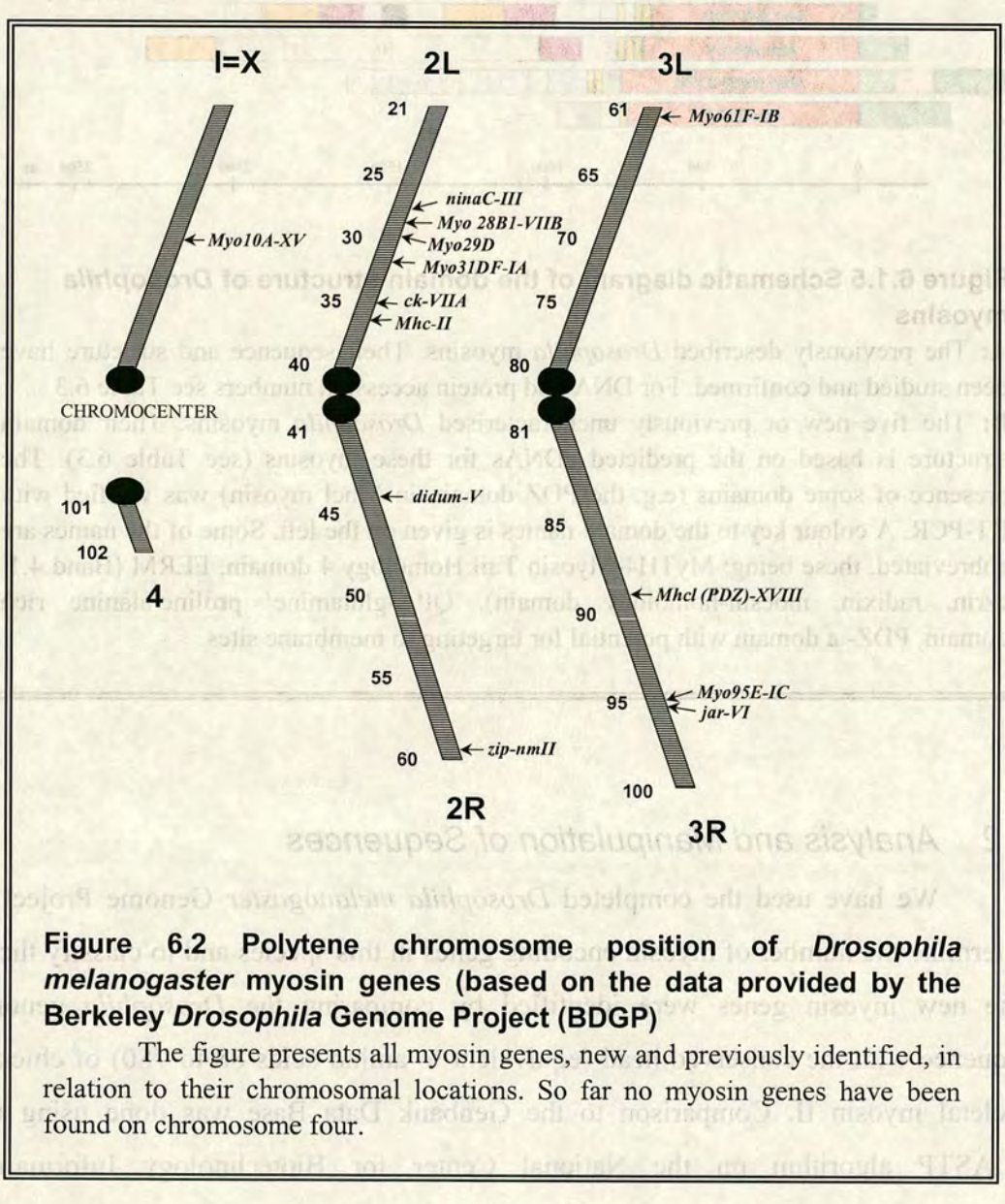


Figure 6.2 Polytene chromosome position of *Drosophila melanogaster* myosin genes (based on the data provided by the Berkeley *Drosophila* Genome Project (BDGP)

The figure presents all myosin genes, new and previously identified, in relation to their chromosomal locations. So far no myosin genes have been found on chromosome four.

6.2.1 Domain Analysis. Multiple Sequence Alignment

The domain structure was predicted with (SMART), Simple Modular Architecture Research Tool server, at http://smart.embl-heidelberg.de/ (Schultz et al. 1998), Pfam HMM database at http://pfam.wustl.edu/hmmsearch.shtml or http://www.sanger.ac.uk/Soft-ware/Pfam/ search.shtml (The Sanger Centre), and the ProFile Scan Server of ISREC (Swiss Institute for Experimental Cancer Research) at http://www.isrec.isb-sib.ch/software/ PFSCAN_form.html. Alignments of the detected domains were performed with CLUSTAL W (Thompson et al. 1994) available from the Gene Jockey II software package distributed by Biosoft or from the WEB based package at http://www2.ebi.ac.uk/clustalw/. Subsequently the sequences were run on WEB based BoxShade Server (http://ludwig-sun1.unil.ch:8080/software/BOX_form.html) and manually adjusted in Microsoft Word 98. Coiled-coil regions were predicted with the Paircoil program developed by Berger and colleagues (Berger et al. 1995) at http://nightingale.lcs.mit.edu/cgi-bin/score.

Primers used in the analysis of *Myo 95E* (Myosin 95E):

Primer name	Sequence in 5' > 3' direction	Length (nt)	Tm (°C)
ST95E (EF1)	TGT TGC TCG CCA GCA CAT G	19	60
1BR	GTG CCG ATT TCC TGC TCC AT	20	62
46F (EF2)	ATG GAG CAG GAA ATC GGC AC	20	62
3FOR (EF3)	CAT CAG CGG CCT TCC TGA AT	20	62
4BAK	TTA GAA CCA CTG CGA TGA CC	20	60
ESR (ER1)	AGG CAG CGC TTA TGT ACT TC	20	60
ESF (EF4)	GAA GTA CAT AAG CGC TGC CT	20	60
P1EF	CGC ACC GTA TAT GAG TTG GA	20	60
9512R	CTA AGA TGA CGA GAG GCA TC	20	60
46R (ER2)	ACA ATT ATC TCC ATG CGG TTC G	22	62
13F	CAT GTC GGA CTT AGC ACA AG	20	60
3UT95 (ER3)	ACG TAG ATG CCT GAA CTA TC	20	58

Primers used in the analysis of *Myo 10A* (Myosin 10A):

Primer name	Sequence in 5' > 3' direction	Length (nt)	Tm (°C)
INF(AF1)	GCA GCA ATG AAT CAA CCG GT	20	60
HYST (AF2)	TGA TCT GGT CTG GTT CGA TC	20	60
510R	AGC AGG GTC ATG TCC TCC A	19	60
10-AR (AR1)	CAT CTT GTA CGG ATT CAC CG	20	60
P15F (AF3)	TCA GTG TCC AGA GGC ATG TG	20	62
MYTH4-10 (AF4)	TGG AGT GGC GTG CCT GGA	18	60
FERM-10 (AR2)	TCC GCA CGC GCA ACT TCC A	19	62
AR3	CTT GCG GAA CTC CTG GAC A	19	60
FERM-12R	GAA GTT GCG CGT GCG GAA G	19	62

Primers used in the analysis of *Mhcl* (Myosin 89B):

Primer name	Sequence in 5' > 3' direction	Length (nt)	Tm (°C)
EVE1 (MF1)	CAA CTT TAT GAA GAA GAG CGC	21	60
NYP2	ATG CTA TTG CGC AAC ACC TTT C	22	62
PDZF-1(new)(MF2)	AAG GCA GCT AGT GAT CAG GC	20	62
PDZ.2R	TCC GCT GTT CCG TAT CAT CT	20	60
589R (MR1)	AGC TCG CAG ATG TCC TCG A	19	60
89DR (MR2)	ACA TGG AGA CAA CCT TCT CG	20	60
CHF (MF3)	TCG CAT AGG ACC CAG CCA G	19	62
89-9F	GGT TCA CAA GTC CTA GAA GC	20	60
EXF (MF4	ATG TGG TCG GAT AAA AGT GCA	21	60
EXR (MR3)	TAT ACG GAC GCA GGC GAT AG	20	62
IN1F (IF1)	TTA CCT CCA TAA ACC TGC GG	20	60
IN1R (IR1)	AAC GAT TCG GAG GTG CAC G	19	60
IN2F (IF2)	GTC GCC GAG CCC GAA GAG	18	62
IN2R (IR2)	CTA GCT CTG CGA AGA TCT CA	20	60
IN3F (IF3)	AAC GTC GCG TTC GCA AGA GG	20	62
IN3R (IR3)	CGT TCA TGG CTG CTA GTA CG	20	62
IN3SP	GAA GAG GAG ACA TCA TCA CC	20	60
389F	AGT GAA CTG GAA GCC AAG CG	20	62

PRPDZF (MF5)	GCT CTC AGA TCG CAT TAT ACA G	22	62
8912R	CTC ACT GCG ATC ACA CTC G	19	60
89-14F	CCG AGA AGT TCA CAC TGG AG	20	62
8914R	GCA GAT CGT TCA TCT CGT TG	20	60
89-15F	GTA GCT CGA AGA ACG TCA AC	20	60
8916R	GTA CTT CTT CCA GCT CTT CG	20	60
PRPDZR (MR4)	TCT TCC AGA TCA CTG ATA GAG	21	60

Primers used in the analysis of *Myo 29D* (Myosin 29D):

Primer name	Sequence in 5' > 3' direction	Length (nt)	Tm (°C)
595F-A (DF1)	ATC CGC ACA ACA TTC TGC AC	20	60
SECF (DF2)	ATG CAT CTT CAT CCA ACG AG	20	58
TRIF (DF3)	AAT CAC AGC TTC AGC CAC AC	20	60
595F-B	ATT GCC ATT CAT CCC GCA GT	20	60
595-R (DR2)	GTG CTT GAA GGC ATC AGT CT	20	60
595F-C (DF4)	GAG ACT GAT GCC TTC AAG CAC	21	62
PHex 29	AGA CAA TTT GCT GAG AAC CCT G	22	62
P29F(DF5)	ACT TTG TGC GCT GCA TCC G	19	60
LASR (DR3)	AGG AAG AGT TGA ACA GAT GGA	21	60

6.3 Results

The BLASTP and BLASTN searches with the conservative chicken skeletal myosin II head, against the completed *Drosophila* Genome sequence, retrieved the previously identified genes and five new myosin genes (Fig. 6.1.5, Table 6.3). We selected a limit of 30-40% identity for a myosin to fall into a given myosin class, and 25-30% was considered as the lower threshold for a protein to be classified as a myosin. The new myosins (Fig. 6.1.5) were named according to their chromosome position (Fig. 6.2). Interestingly half of the myosin genes are located between polytene bands 27F-36A, thus forming a myosin hot spot on the left arm of the 2 chromosome.

In the cases of *Myo 95E, Myo10A, Mhcl and Myo29D* we did a detailed analysis of the molecular structure of the genes and the transcripts they produce. The main reason for this was the predicted sequences for these genes encode proteins that produced low homology scores to other myosins. They were obviously myosins, containing all the conserved sequences and structural parts defining them as myosins and at the same time showing no more than 15-29 % identity to other myosins. This implied either incorrectly predicted genes or incorrectly predicted splicing of transcripts. Open reading frames (ORF), 5'UTRs (untranslated regions), 3'UTRs, and the presence or absence of given motifs were tested by RT-PCR and subsequent sequencing of the products obtained. In these experiments we employed an ovarian Uni-ZAP XR® library produced in our lab, along with cDNAs produced by reverse transcription RNA from larvae and adult flies RNA (see materials and methods). As a result we determined a number of new myosin sequences and submitted them to the MEDLINE Database. The accession numbers for these are: for *Myo95E* (AF454350, AF454351 and AF454352), For Myo10A (AF454346, presents a part of the first three exons including the 5'UTR), *Mhcl* (AF454347, presents a part of the first four exons including the sequence encoding the PDZ domain), *Myo29D* (AF454348 and AF454349). *Myo28B* was not subjected to detailed analysis because it was found to be almost identical to the other myosin VII (*crinkled*) from *Drosophila* at both the DNA and protein level.

Table 6.3 A list of *Drosophila melanogaster* myosins

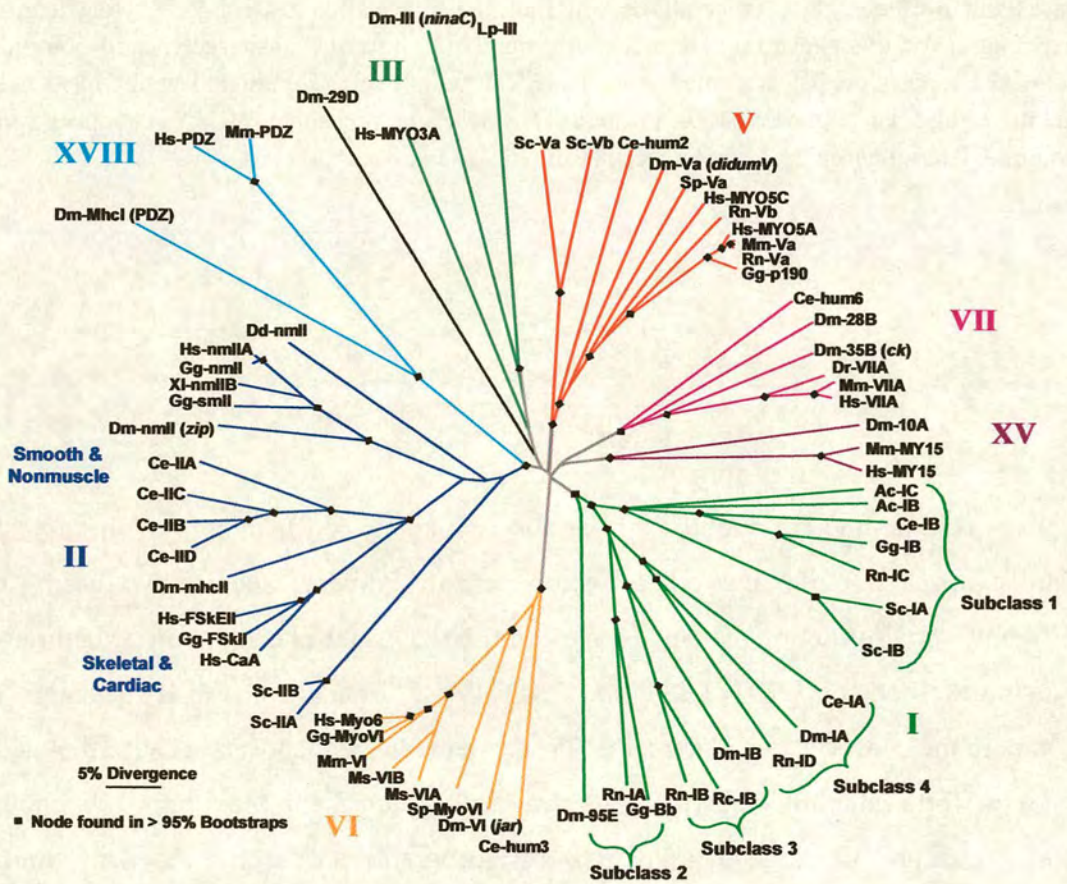
Gene/ Protein	Class	Chromo-some position	DNA Length (bp)	mRNA Length (bp)	Protein Length (aa)	EST
Myo31DF **(CG7438)*** **Myosin IA**	I	2L (31E3-6)	AE003628 (15084)	U07595 (3886)	AAA19590, AAF52966 (1011)	LP03284.5, LP04189.5,
Myo61F **(CG9155)*** **Myosin IB**	I	3L (61F6-7)	AE003471, AC005847 (5057)	U07596 (3382)	AAA19591 (1026)	GH13657.5, LP05759.5, GH13670.5, LP08690.5, GH25605.5, GH23689.5, CK01057.5
Myo95E **(CG5501)** **Myosin IC**	I	3R (95E1-4)	AE003746 (4604)	AF454350 (5285) AF454351 (5225) AF454352 (5301)	AF454350 (1278) AF454351 (1258)	GH25580.5, GH25580.3, HL05373.5
zip (zipper) **(CG15792)*** **Nonmuscle myosin-II**	II	2R (60E7-8)	AE003465 (21711)	U35816 (6734)	AAF47311, AAB09048 (2056)	GM04963.5, SD07905.5, GH27070.5
Mhc **(CG17927)*** **Myosin heavy chain**	II	2L (36A8-9)	M61229, AE003652, AC005119 (22663)	M61229 (5889)	AAA28686, AAA28687 (1962)	GH21445.5, GH06021.5, GH02291.5, LP07131.3, LD16079.5, HL04385.5, GH11211.5, LD35610.3, HL03720.5, GM03715.5, GM03876.5
ninaC **(CG5125)*** **Nina C protein**	III	2L (27F5-6)	AE003617, J03131, (7220)	M20230 (4809) M20230 (3643)	AAA28718 (1501) AAA28719 (1135)	LP02603.5, LP02603.3
didum **(CG2146)*** **Myosin V**	V	2R (43D1-3)	AC004280, AE003841 (8217)	AF003826 (6214)	AAF59241, AC99496 (1792)	GH04445.5, GH04445.3
jar (jaguar) **(CG5695)*** **Myosin VI**	VI	3R (95F1-2)	AE003747 (6312)	X67077 (4280)	AAF56269, CAA47462 (1253)	HL03149.5, GH09735.5
ck (crinkled) **(CG7595)*** **Myosin VIIA**	VII	2L (35C1)	AE003646 (12392)	AE003646 (7030)	AAF53435, AAF44915 (2167)	LD14917.5, LD14917.3
Myo28B1 **(CG6976)** **Myosin VIIB**	VII	2L (28B3-C1)	AE003618, AC005834 (10519)	AF233269 (6590)	AAF52536, AAF34810 (2129)	GH25551.5, GH25551.3
Myo10A **(CG2174)** **Myosin XV**	XV	X (10A1)	AE003484 (10235)	AE003484, AF454346 (7275)	AAF47980, Q9VZ48 (2424) (2333)	LP03318.5
Mhcl **(CG10218)** **Myosin heavy chain like, (PDZ) Myosin**	XVIII	3R (89B7)	AE003711 (21620)	AE003711 AF454347 AY051503	AAK92927, AF454347, AAF55271, AAF55272 (6603) (6603) (6420)	GH03004.5, GM10420.5, GH04935.5, LP04491.5, GH15471.5, GH15471.3
Myo29D **(CG10595)** **Myosin**	?	2L (29D1)	AE003621 (5732)	AF454348 (4402) AF454349 (2530)	AF454348 (1313) AF454349 (689) AAF52683, Q9VLK6,AF405293 AAK97502	LD47348.5, LD47348.3, LP07160.5, AI124339

Table 6.3 A list of *Drosophila melanogaster* myosins

Myosins denoted with * have been previously identified and characterised. The others resulted from analysis of the published genomic sequence for *Drosophila melanogaster* and molecular studies undertaken in this study. Their classification is based on domain and phylogenetic analysis of the predicted mRNAs for these myosins. In several cases the presence or absence of certain motifs was verified by RT-PCR (for detailed explanation of this see text). The accession numbers show the most recent and complete sequence reports on the presented myosins. ESTs for all of the predicted genes have been identified and partially sequenced. In the case of *Mhcl* the presented mRNA sequences were combined and the longest three transcripts shown.

To examine the evolutionary relationships between members of the myosin family in *Drosophila* and other phylogenetically diverse species we used two different phylogenetic methods. We applied Distant-matrix, and Maximum-Parsimony methods (PROTDIST and PROTPARS from the PHYLIP package) to compare the conserved head domains. These methods were chosen because they tend to outperform other methods (i.e., lower variance), such as the Maximum Likelihood, when dealing with large datasets. The two programs produced trees with similar topology (see the unrooted consensus tree in Fig. 6.3). Multiple sequence alignments were performed with CLUSTAL W, without corrections for gaps or multiple substitutions. Excluding the positions with gaps would have omitted a significant proportion of the data, a problem that occurs when large amounts of input sequences are dealt with. CLUSTAL W (GCG software package) is provided by the Human Genome Mapping Project Resource Centre, Cambridge at http://www.hgmp.mrc.ac.uk/ (Thompson et al. 1994). The reliability of the tree structure was checked by bootstrapping (1000 trials) and re-ordering the alignments randomly (bootstrapping was performed with SEQBOOT from the PHYLIP package). The tests produced trees with similar branching order. A consensus tree was produced by the CONSENSE program of the PHYLIP package and graphically drawn with the TREEVIEW program (Page 1996), then transferred to and manipulated with PowerPoint.

Figure 4 Phylogenetic analysis of the myosin superfamily in *Drosophila melanogaster*



Abbrevation full terms:

Ac	*Acanthamoeba castellanii*	Lp	*Limulus polyphemus*	Bb	Brush border myosin I
Ce	*Caenorhabdites elegans*	Mm	*Mus musculus*	Ca	Cardiac alpha myosin II
Dd	*Dictyostelium discoideum*	Ms	*Morone saxatilis*	FSk	Fast skeletal muscle myosin II
Dm	*Drosophila melanogaster*	Rc	*Rana catesbeiana*	mhc	Myosin heavy chain
Gg	*Gallus gallus*	Rn	*Rattus norvegicus*	nm	Non-muscle myosin II
Hs	*Homo sapiens*	Sc	*Saccharomyces cerevisiae*	PDZ	PDZ containing myosin XVIII
		Xl	*Xenopus laevis*	sm	Smooth muscle myosin II

Figure 4 Phylogenetic analysis of the myosin superfamily in *Drosophila melanogaster*

The head domains of *Drosophila* myosins (equivalent to amino acids 88 to 780 of chicken skeletal myosin) were compared to selected myosins representing phylogenetically diverse organisms. The alignment of 67 myosins (performed with the CLUSTAL W package) was used to generate data for an unrooted tree to scale, based on sequence divergence (for details see in the text). Each separate myosin class is defined by the first node represented in > 95% of bootstrap trials (starting from the tree centre) and is colour coded. For simplicity only the *Drosophila* myosins are presented with the names of the genes: *Mhc(Myosin heavy chain), zip (zipper), jar (jaguar), ck (crinkled)* and *Mhcl (Myosin heavy chain like)*. The accession numbers for the myosins used in the alignment are given below. These are produced for the protein sequences except in the case of Dm95E (*Myo95E* gene) and Dm29D (*Myo29D gene*), where the cDNA accession numbers are given:
MYOSIN CLASS I: Ac-IB (AAA27708), **Ac-IC** (AAA27707), **ScMyo3-IA** (AAB34124), **ScMyo5-IB** (CAA89745), **Cehum5-IA** (CAA53244), **CehumI-IB***, **Dm31D-IA** (AAA19590), **Dm61F-IB** (AAA19591), **Dm95E** (AF454350), **Rc-IB** (AAA57192), **Gg-Bb** (AAB38373), **Gg-IB** (CAA49850), **Rnmyr1-IA** (CAA48287), **Rnmyr2-IB** (CAA52807), **Rnmyr3-IC** (CAA52815), **Rnmyr4-ID** (CAA50871). **MYOSIN CLASS II: ScMYO1-IIA** (CAA37894), **ScMYS1-IIB** (P08964), **Dd-nmII** (AAA33227), **Cemyo3-IIA** (CAA30856), **Cemyo4-IIB** (AAA28124), **Cemyo2-IIC** (CAA30855), **Cemyo1-IID** (CAA30854), **Dm-mhcII** (AAA28686), **Dmzip-nmII** (AAF47311), **Xl-nmIIb** (AAA49915), **Gg-FSkII** (P13538), **Gg-nmII** (AAA48974), **Gg-smII** (CAA29793), **Hs-FSkEII** (CAA32167), **Hs-CaA** (BAA00791), **HsnmIIA** (AAA59888). **MYOSIN CLASS III: Lp-III** (AAC16332), **DmNinaC-III** (AAA28718), **HsMYO3A** (AAF70861). **MYOSIN CLASS V: ScMyo2p-Va** (AAA34810), **ScMyo4p-Vb** (AAC37409), **Cehum2** (AAA97926), **SpMyo-Va** (AAF78910), **Dmdidum-Va** (AAC99496), **Ggp190-Va** (CAA77782), **Mmdilute-Va** (CAA40651), **Rnmyo5-Va** (BAA88350), **Rnmyr6-Vb** (AAB38840), **HsMYO5A** (AAD00702), **HsMYO5C** (AAF78783). **MYOSIN CLASS VI: SpMyo-VI** (AAF72176), **Cehum3** (AAC67447), **Dmjar-VI** (CAA47462), **Ms-VIA** (AAD52005), **Ms-VIB** (AAD52006), **GgMyoVI** (CAB96536), **MmWaltzerVI** (AAB00194), **HsMYO6** (AAC51654). **MYOSIN CLASS VII: Cehum6-VII** (AAB37988), **Dmck-IIA** (AAF53435), **Dm28B-VIIB** (AAF52536), **DrMyo-VIIA** (CAC05418), **Mmshaker-VIIA** (AAB40708), **HsUsherIb-VIIA** (AAB03679). **MYOSIN CLASS XV: Dm10A-XV** (AAF47980), **Mm-Myo15** (AAF05904), **HsMYO15** (AAF05903). **MYOSIN CLASS XVIII: Dm89B-PDZ** (AAF55271), **Mm-PDZ** (BAA93660), **Hs-PDZ** (BAA13206)**. **MYOSIN CLASS ?: Dm29D** (AF454348).

* This is a joint sequence from the first 72 amino acid of T21544 and from the 11th amino acid to the end of AAA97925.

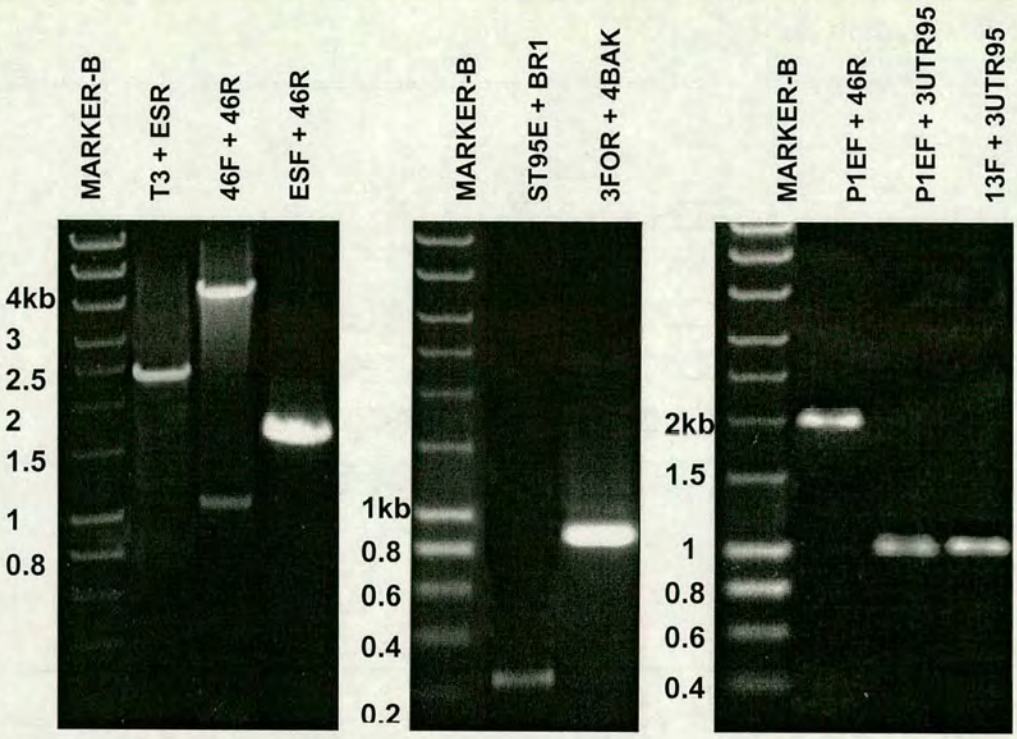
** The sequence for the human PDZ-containing myosin is truncated. This does not seem to affect its evolutionary position in the phylogenetic tree.

The protein sequences for the new *Drosophila* myosins are theoretical predictions. Myosins are large multi-exon genes and are difficult to assemble with 100% accuracy from sequence data. There are also various isoforms of some myosins, which can lead to some misalignments; hence it is unlikely that the tree shown perfectly reflects the evolution of *Drosophila* myosins.

It is possible that some of the new *Drosophila* myosins could be pseudogenes. However, we found that the probability for this was low. Pseudogenes generally lack introns and are not transcribed into mRNA. We identified ESTs (expressed sequence tags) for all the myosin genes we predicted (the accession numbers for these are given in Table 6.3) which confirms their *in vivo* expression. The new myosins are described in details below.

6.3.1 Myo95E (Myosin IC)

This myosin was not found during the initial searches of the fly genome. It was subsequently identified in the AE003746 genome scaffolding. The predicted sequence for this gene is unusually short, resulting in only 59 amino acid protein (accession number AAF56246). Detailed analysis showed that the gene was not predicted correctly by the Genome Project. The sequence for *Myo95E* was manually assembled, taking into account the reported sequence for this gene, along with the identified ESTs (Table 6.3), as well as the homology of the translated genomic DNA to other myosins (preserving the exon-intron spacing). To test our theoretical predictions we used two splice-site prediction programs: The Neural Network at http://www.fruitfly.org/seq_tools/splice.html and the GENSCAN Server at MIT: http://genes.mit.edu/GENSCAN.html. Using the artificially assembled sequence we were able to design primers, create a PCR product and hence sequence the cDNA for *Myo95E* (Fig. 6.3.1 and Appendix II). The resulting sequence differed from the predicted sequence showing the presence of an unusually long exon 3 and variations in exon 4. Analysis of the available ESTs and RT-PCR analysis showed that the gene produces at least three different transcripts expressed during oogenesis, larval and adult stages (Fig.6.3.1B).



PCR Primers	Expected product	Amplified/Sequenced product
46F + ESR	2315 bp	2315 bp
46F + 46R	3849 bp	3849 bp*
ESF + 46R	1554 bp	1554 bp
ST95E + BR1	279 bp	279 bp
3FOR + 4BAK	914 bp	914 bp
P1EF + 46R	1085 bp	1085 bp
P1EF + 3UTR95	2084 bp	2084 bp
13F + 3UTR95	1021 bp	1021 bp

Figure 6.3.1 RT-PCR analysis of *Myo95E* transcripts

mRNA was prepared from different developmental stages and reverse transcribed. An ovarian cDNA library was also used as a source of mRNA. Gene-specific primers were used to generate overlapping PCR products. The latter were analysed by sequencing and assembled into full length mRNA using the program GeneJockey. New primers, based on this sequence, were developed to analyse the exon structure of *Myo95E*. The available ESTs for the gene were also considered (Fig.6.3.1B).

*The amplification with 46F and 46R primers produced two bands. Only the large 3.8 kb band is genuine while the small 1.1 kb band is a product of mispriming.

Figure 6.3.1B Molecular organisation of *Myo95E, Myo10A, Mhcl, and Myo29D* myosin genes

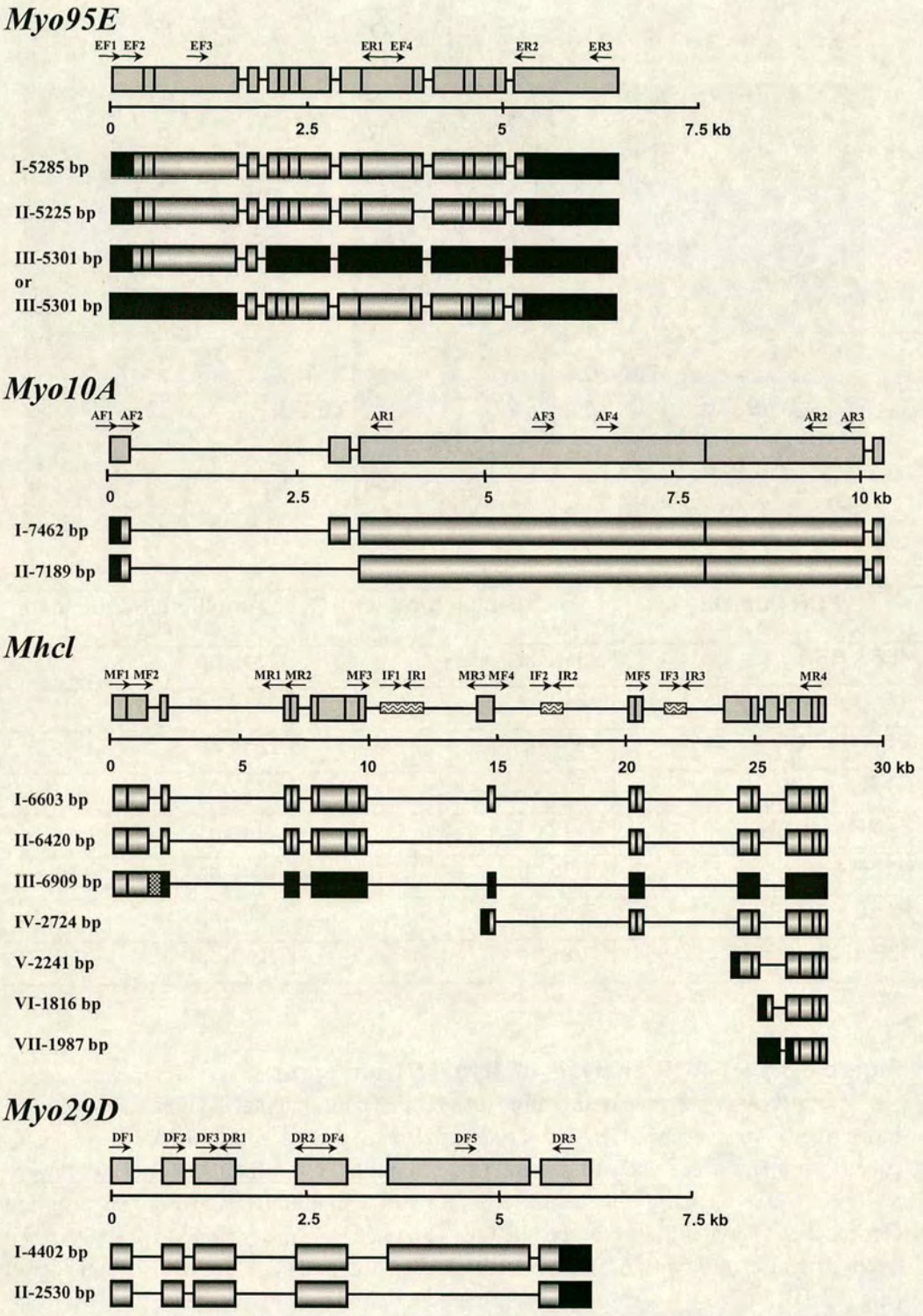


Figure 6.3.1B Molecular organisation of *Myo95E, Myo10A, Mhcl, and Myo29D* myosin genes

The exon structure of the myosin genes is represented by grey boxes. Arrows indicate the primers used in the RT-PCR analysis. The primer sequences are given in the materials and methods. The first letter of the primer name is specific for each myosin, the second letter indicates whether the primer is in the forward (F) or reverse (R) orientation. The length of the predicted mRNAs is noted on the left side of the transcript maps. Shaded boxes represent open reading frames (ORFs); solid boxes represent untranslated sequences (5'UTR or 3'UTR); dotted boxes represent intron sequence in the spliced mRNA; wave pattern boxes represent exons from genes located within the myosin gene (the first letter of the primer name for this genes is I from Intron). The gene depicted by IF1 and IR1 primers is at least a 3 exon gene, but for simplicity it is presented with a single box here (for more information see text).

Two of the transcripts (I-AF454350 and II-AF454351), which represent 5-10% of the total amount of *Myo95E* transcripts, translate into two protein isoforms: of 1278 and 1258 amino acids, respectively. Transcript One (5285 bp) comprises all 16 exons of the gene. Transcript Two (5225 bp) lacks exon 12. The major third transcript (III-AF454352) has a longer 4th exon (a downstream extension of 16 bp, GTG CAC ATT ACC CAT T). This shifts the open reading frame and produces a STOP codon TGA in exon 5. The third transcript (5301 bp) encodes two putative truncated proteins. The first is a 464 aa sequence containing only the GESGAGKT conserved region from the P-loop of the myosin head domain. It does not contain the Switch-1 (NxxSSR) and Switch-2 (DxxGxE) regions which together with the P-loop have been implicating in the regulation of ATP hydrolysis. The second is a 864 aa putative protein, containing a large part of the head domain, two IQ motifs and a myosin tail domain. The head of this isoform contains only the Switch-2 (DFYGFE) conserved sequence, which prevents it from converting ATP, probably rendering the isoform inactive. Homology searches with the protein sequence for the head domain showed 33% identity (53% similarity) to vertebrate brush border class I myosins. Three ESTs were found for *Myo95E* . A search with them revealed up to 37% identity to vertebrate brush border myosins and less to other myosin classes. Analysis of Myosin 95E with domain scanning programs confirmed its structural similarity to class I myosins (Fig. 6.1.5B). It was found to contain 2 IQ domains (928-974 aa) the

second being poorly conserved (Fig. 6.3.1C), and a Basic Tail domain (974-1278 aa for isoform I and 974-1258 aa for isoform II) (Fig.6.1.5, Table 6.3). The latter is thought to be involved in membrane binding. Recent studies have shown it is also able to bind to actin filaments (Lee et al. 1999) and (Liu et al. 2000). This changes the number of class I myosins in *Drosophila* to three, hence Myosin 95E was renamed Myosin IC. The *Drosophila* Myosin IC does differ from other class I myosins. The unusually long exon 3 results in a 281 amino acid insertion into the head domain. This insertion contains a partial AAA domain, a conserved region that contains an ATP-binding site. So far no other myosins from class I have been identified that contain such an insertion.

Figure 5 Alignment of the IQ motifs for the five newly identified myosins from *Drosophila melanogaster:* Myosin IC *(Myo95E),* Myosin VIIB *(Myo28B),* Myosin XV *(Myo10A),* Myosin89B *(Mhcl), and* Myosin 29D*(Myo29D)*

The consensus sequence is shown below the alignment and the highly conservative positions are shaded. Residues in bold indicate similarity to the consensus. Myosin IC has two different variants (A and B) of its second IQ motif, a result of alternative splicing.

6.3.2 Myo28B *(Myosin VIIB)*

The amino acid sequence of the Myosin 28B head showed 61% identity (74% similarity) to *ck-Drosophila* myosin VIIA. It also exhibited a very high identity of 58% (72% similarity) to Myosin VIIA from zebrafish and other class VII myosins. Analysis of Myosin 28B revealed that it has four IQ motifs (753-845 aa), the third being poorly conserved (Figs 6.1.5 and 6.3.1C). Two MyTH4 (Myosin Tail Homology 4 domain) (1070-1246 aa and 1681-1826 aa), two FERM (1246-1454 aa and 1826-2039 aa), and one SH3 domain (1561-1626 aa) were identified (Fig. 6.3.2A, B).

The function of MyTH4 domains is unknown. The FERM domain, (the name stands for Band 4.1, ezrin, radixin, moesin-homology) is believed to be involved in linking cytoskeletal proteins to the membrane, as well as in dimerisation. Talin, merlin and philopodin are other major members of the FERM superfamily. These deserve mentioning, especially the talin, because the FERM domain exhibits highest homology to the FERM domain in talins (Fig. 6.3.3) and less homology to FERM domains from other members of the FERM family.

The SH3 domain has been identified in many proteins involved in signal transduction. It is believed that SH3 domains mediate protein-protein interactions by binding to proline-rich domains. Other myosins such as IV, X and XV also contain this motif. A short coiled-coil domain (849-908 aa) was predicted by the Paircoil program (Fig. 6.3.4).

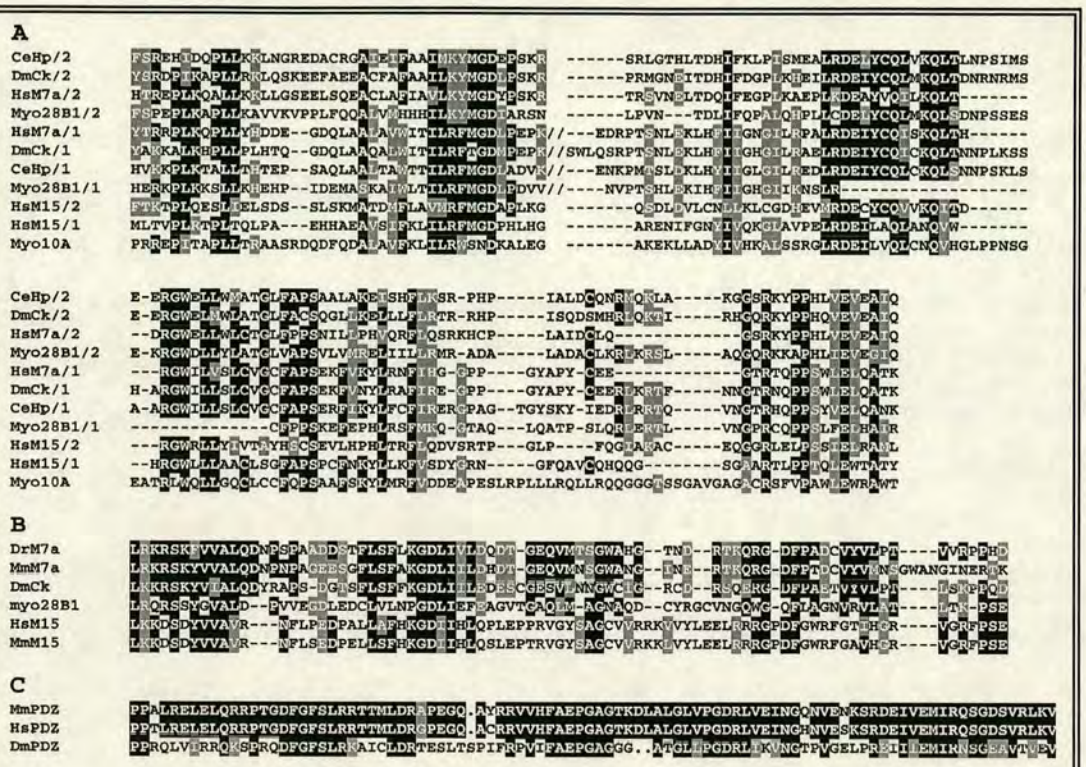
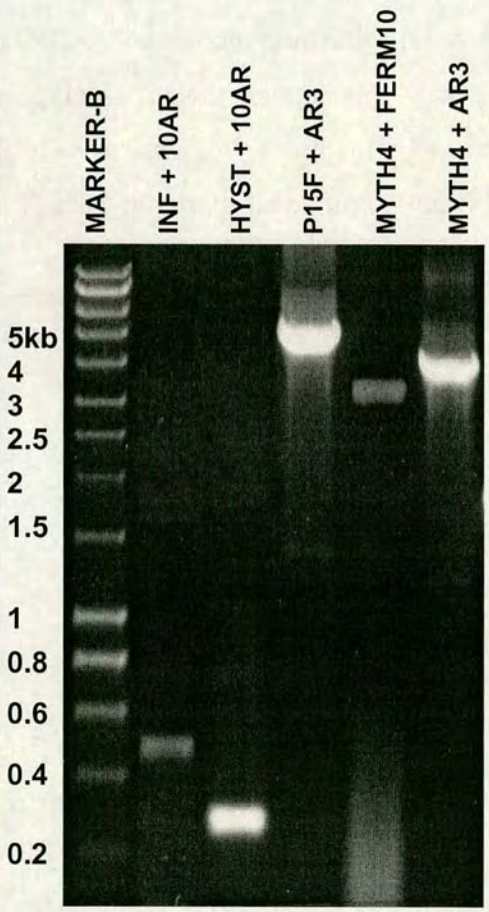


Figure 6.3.2 CLUSTAL W alignment of the predicted MyTH4 (A), SH3 (B), and PDZ (C) domains in the tails of the newly identified *Drosophila* myosins

The sequences were shaded on the BoxShade server and manipulated with Microsoft Word 98. The parameters set on BoxShade are as follows: output format: RTF New; Fraction of sequences: 0.5; Input sequence format: MSF (multi sequence format this is the output format of CLUSTAL W). Identical amino acids are shaded in black, similar residues are shaded in grey. Myosin VII has two MyTH4 domains. The first MyTH4 domain in HsM7a, DmCk, CeHp, and Myo28B1 contain an insert of 61-88 aa (regions 1057-1144; 1048-1135; 971-1055; 1107-1168, respectively) that was not included in the alignment. Accession numbers and Abbreviation full terms: Ce-*Caenorhabditis elegans* (Myosin VII-T25888), Hp-hypothetical protein, Dm-*Drosophila melanogaster* (ck, VIIA-AAF44915, Myo28B-AAF52536, PDZ Myosin-AE003711), ck-crinkled, Hs-*Homo sapiens* (M7A-Q13402; Myosin XV-A59266, PDZ Myosin-AL080245 and Z98949), Dr- *Danio rerio* (Myosin VIIA- CAC05418), Mm-*Mus musculus* (Myosin VIIA-P97479, Myosin XV-A59295, PDZ Myosin-BAA93660).

6.3.3 Myo10A *(Myosin XV)*

An RT-PCR analysis was used to verify the exon composition and exon length of *Myo10A* transcripts (See Appendix III). It was found that there are at least two transcripts which are expressed from the larval stage onwards (Fig 6.3.1C). The longer transcript (7462 bp) consists of all the previously known 5 exons and represents no more than 5% of the total amount of mRNA for this myosin. The shorter transcript (7189 bp) lacks exon 2 and is expressed abundantly. The two transcripts translate into two protein isoforms: 2424 and 2333 amino acids respectively.



PCR Primers	Expected product	Amplified/ Sequenced product
INF + 10AR	514 bp	514 bp
HYST + 10AR	327 bp	327 bp
P15F + AR3	4703 bp	4703 bp
MYTH4 + FERM10	3194 bp	3194 bp
MYTH4 + AR3	3746 bp	3746 bp

Figure 6.3.3 RT-PCR analysis of *Myo10E* transcripts

mRNA was prepared from different developmental stages and reverse transcribed. Primers specific for *Myo10A* were used to generate PCR products. The latter were analysed by sequencing and assembled into full length mRNA using GeneJockey.

A BLAST search with the conserved head domain showed significant, 47% identity (64% similarity), to the mouse and human myosin XV. *Myo10A* is also related to human and mouse myosin VIIA with an identity of 42% (59% similarity). A specific N-terminal domain was identified in Myosin 10A (1-149 aa). The latter showed low similarity to the characteristic N-terminal domain found in other class XV myosins. The shorter protein isoform lacks this N terminal domain. In the neck region three IQ domains (841-910 aa) were identified (Fig. 6.3.1C). Immediately after the IQ motifs there is a short coiled-coil region (919-946aa) (Fig. 6.3.4). Analysis of the tail revealed the presence of one MyTH4 domain (1014-1173 aa) (Fig. 6.3.2), a QPA-rich domain, a Proline rich domain (the borders of QPA and the Proline rich domain were not clearly defined), a short transmembrane motif (2194aa) (Fig.6.1.5), and a FERM domain (2220-2424aa). The latter showed a very high identity of 40% (59% similarity) to the first and 13% identity (30% similarity) to the second FERM domain from mouse myosin XV and a limited similarity of 38% (17% identity) to Talin itself (Fig. 6.3.3).

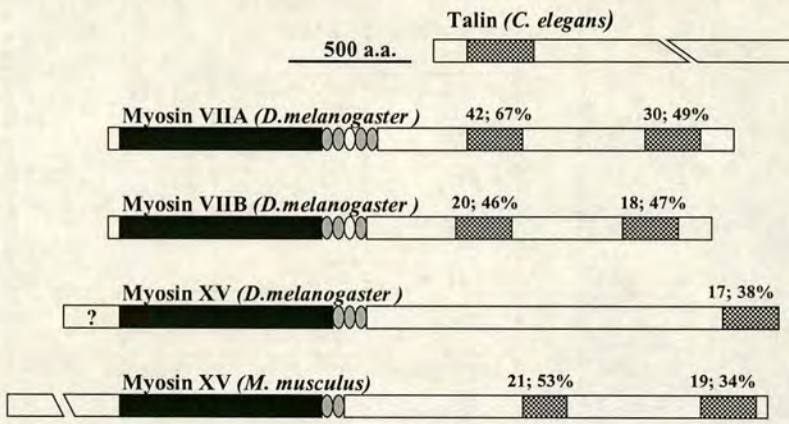


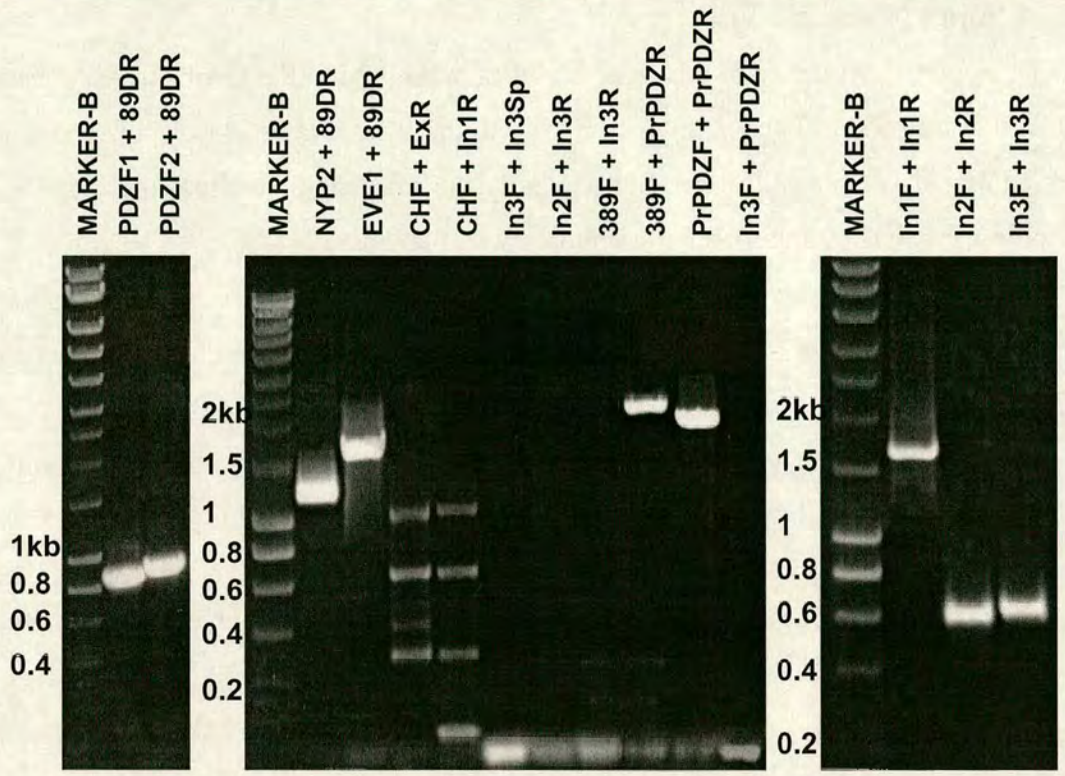
Figure 6.3.3 Sequence comparison of the FERM domains found in *Drosophila* Myosin VII and XV

These were aligned separately to the Talin domain from *C. elegans* Talin protein. Multiple alignment was not appropriate in this case because the similarity between FERM 1 and FERM 2 to the Talin domain (from the Talin protein) is very low and results in unordered alignment. The comparison was performed by CLUSTAL W (http://www2.ebi.ac.uk/clustalw/). In addition the FERM domains from *M. musculus* Myosin XV were included in the diagram. In general FERM1 motifs from Myosin VII show ~12% to 21% similarity to proteins from the FERM family and ~16% to the band 4.1 protein. FERM2 domains (Myosin VII), show a higher similarity of ~17-18% to band 4.1 protein and ~18-20% to proteins from the FERM family. FERM domains from Myosin XV show less similarity of 9%-17% to FERM proteins and ~14% to the band 4.1 protein. Accession numbers are: *C. elegans* (Talin)-AAA74747, *D. melanogaster* (Myosin VIIA)- CAC05418, (Myosin VIIB)-AAF52536, (Myosin 10A-XV)-AAF47980, *M. musculus* (myosin XV)-AAF05904.

6.3.4 Mhcl *(Myosin XVIII)*

A partial sequence of this gene has been submitted to the Genbank Data Base by Biru (Biru 1999). It shows high similarity to the Mhc II type of myosins and was termed *Myosin heavy chain like*. Subsequently the full length (without the sequence encoding for PDZ domain) for this myosin was published by Celera (Adams et al. 2000). RT-PCR analysis we have undertaken shows that *Mhcl* has a very complex structure and expresses multiple transcripts throughout the *Drosophila* life cycle (Fig 6.3.1B).

Data from the available ESTs for *Mhcl*, PCR analysis and a Promoter predicting program (http://www.fruitfly.org/seq_tools/promoter.html) suggest that the seven identified transcripts are transcribed from four putative promoters (See Appendix IV). Transcripts I (6603 bp) and III (6909 bp) are highly expressed. The other five identified transcripts are expressed at comparatively lower levels. Transcripts I and III are composed of 18 exons (the gene comprises 19 exons), lacking exon 15. The difference between them is that in transcript III the second intron is not spliced out which introduces a STOP codon in the open reading frame. One unusual feature of *Mhcl* is the fact it contains three other genes within it (Fig. 6.3.1B). The exons of these genes are not found in combination with any of the exons of the myosin gene. Importantly there are ESTs for each of the three genes. One of the genes has been previously identified as phosphatidylserine-specific phospholipase A1 (CG4979). Its is located in the 9^{th} intron and its reading frame is in that opposite direction to the *Mhcl* reading frame (EST: GH15759). The other two genes have open reading frames in the same direction as the main (*Mhcl*) gene. The gene in intron10 (ESTs: RE41368, RE44374) is a novel gene and does not show any significant homology to previously characterised genes. The gene in intron 12 (ESTs: LP08646, LP05315) shows a limited identity of 35% (46% similarity) to bovine synaptojanine 1 protein (SYNAPTIC INOSITOL-1,4,5-TRISPHOSPHATE 5-PHOSPHATASE 1).



PCR Primers	Expected product	Amplified/Sequenced product
PDZF1 + 89DR	848 bp	848 bp
PDZF2 + 89DR	893 bp	893 bp
NYP2 + 89DR	1264 bp	1264 bp
EVE1 + 89DR	1852 bp	1852 bp
CHF + ExR	/ 202 bp ?	300, 400, 600 and 1.0 kb *
CHF + In1R	/ ?	50,300, 600, and 1.0 kb *
In3F + In3Sp	/ (1841 bp) ?	/ (1841 bp) ?
In2F + In3R	/	/
389F + In3R	/	/
389F + PrPDZR	2312 bp	2312 bp
PrPDZF + PrPDZR	2281 bp	2281 bp
In3F + PrPDZR	/	/
In1F + In1R	1585 bp	1585 bp
In2F + In2R	611 bp	611 bp
In3F + In3R	652 bp	652 bp

Figure 6.3.1 RT-PCR analysis of *Mhcl* transcripts

mRNA was prepared from different developmental stages and reverse transcribed. An ovarian cDNA library was also used as a source of mRNA. Gene-specific primers were used to generate PCR products. The available ESTs for this gene were extensively analysed and some of them sequenced to determine the intron-exon structure of the gene. These date was used to assembled the full length Myo89B mRNA. *Initial PCR experiment using CHF primer produced a multiple pattern on agarose gel because of mispriming. Subsequent PCRs produced two bands as expected, a 532 bp band and a 349 bp band that indicate the lacking of exon 10 in the produced transcripts (data not shown).

Mhcl transcripts translate into a set of protein isoforms: I-2200 aa, II-2139 aa, III-439 aa, IV-899 aa, V-730 aa, VI-512 aa and VII-479 aa. Only the isoforms produced by transcripts I and II contain the conserved motor domain. BLAST analysis showed that Myosin heavy chain-like is most closely related to mouse PDZ-Myosin (41% identity) and human PDZ-Myosin (40% identity). It showed a limited identity of 28-29% to smooth muscle and non-muscle myosins and less than 25% to skeletal muscle myosin. BLAST searches revealed that the tail domain shows the highest similarity, 32%, to PDZ-containing myosins and less than 25% similarity to other types of myosins. Interestingly this myosin was not predicted to contain a PDZ domain. We used the DNA sequence encoding the mouse PDZ domain to search the genomic sequence (AE003711) surrounding *Mhcl* (approximately 10kb in each direction). This detected a PDZ domain 4810bp upstream of the predicted start for *Mhcl* mRNA (the domain was found in the borders of the predicted genomic DNA for this gene but has not been previously included in its open reading frame (ORF)). The presence of the PDZ domain was verified by RT-PCR and the sequence submitted to the NCBI database (AF454347). A sequence alignment of the PDZ domains from *Drosophila* mouse and human is given on Fig. 6.3.2. Interestingly transcript III translates into a truncated protein containing only this PDZ domain (349-429 aa). PDZ domains are known to bind C-terminal or internal (non-C-terminal) polypeptides. Two IQ domains were found in Mhcl (1379-1428 aa) (Fig. 6.3.1C), although the second differed slightly from the consensus sequence. The Paircoil program predicted two coiled-coil domains, at 1439-1549 aa and 1616-2121 aa (Fig. 6.3.4).

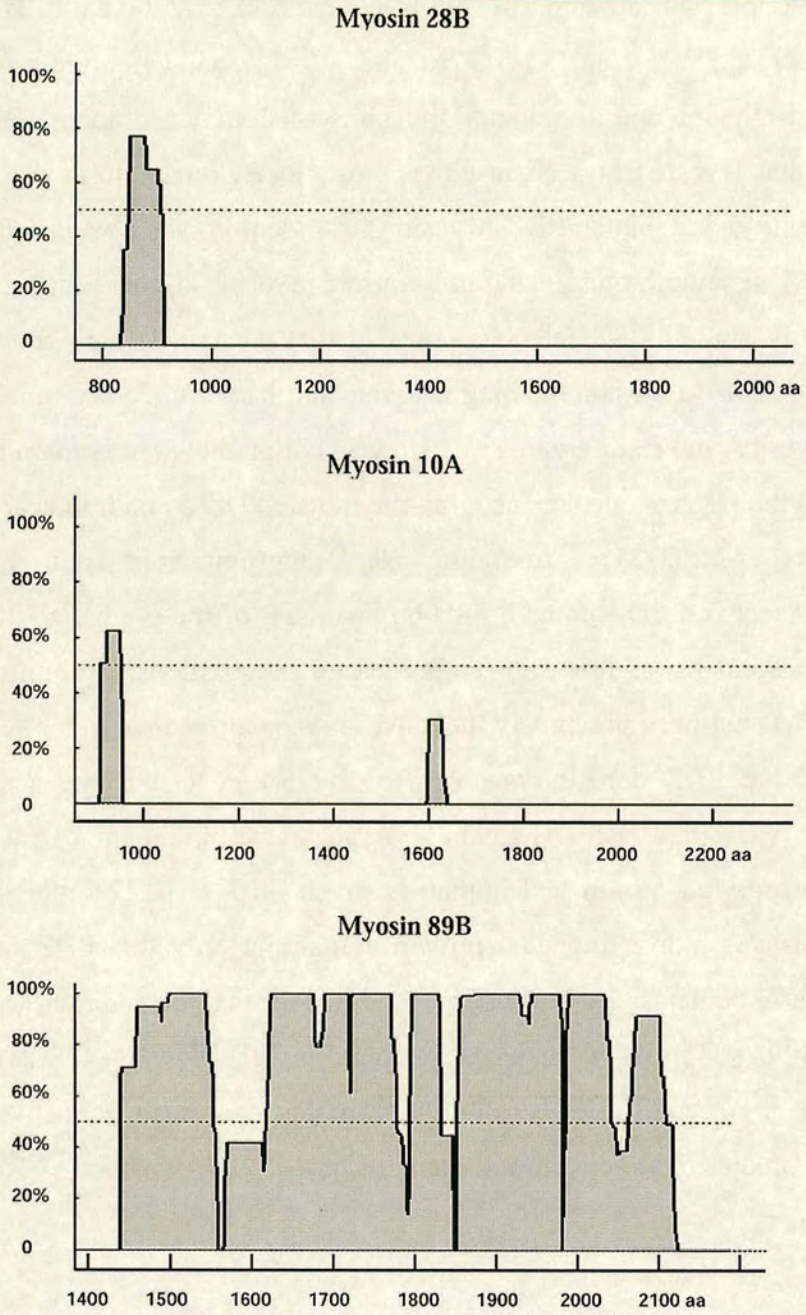


Figure 6.3.4 Predicted coiled-coil structures in the tails of the newly identified *Drosophila* myosins

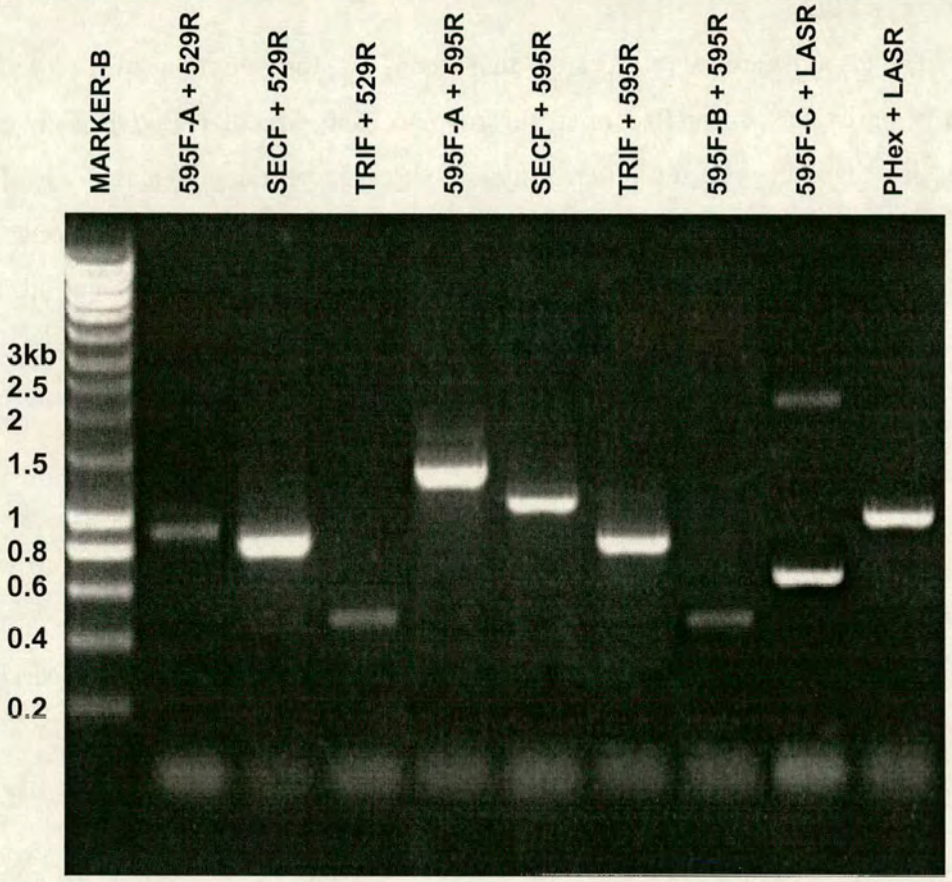
Only three of the five myosins are likely to form coiled-coils, these being Myosin 28B-VIIB, Myosin 10A-XV, and Mhcl-PDZ. Both myosin 28B-VIIB and 10A-XV produced scores close to the lower threshold. Whether they form coiled-coils or not can only be determined experimentally.

6.3.5 *Myo29D*

Database searches suggested that the first three exons of *Myo29D* are probably part of a gene adjacent to the myosin. Subsequent RT-PCR showed that they splice together with the other 3 myosin exons to produce at least two different transcripts (see Appendix V). Transcript I (4402 bp) contains all 6 exons, while transcript II (2530 bp) lacks exon 5. Transcript I and II are present during all stages of *Drosophila* development, with transcript I being expressed at higher levels. Transcript II produces a truncated form of the protein lacking most of the motor head domain.

Analysis of the conserved head domain revealed that Myosin 29D is not similar to any of the known classes of myosins. It showed 29% identity (45% similarity), to myosin VII, X, and V from different species, which is sufficient to be considered as a myosin, but not enough to be related to a given class. A search with the available ESTs for this myosin showed 36 to 39% identity to Myosin V from different species and 31-36% to vertebrate myosin heavy chains (Myosin II). Motif search programs detected in Myosin 29D a specific N-terminal extension (1-338 aa), two IQ motifs (1089-1136 aa), the second being poorly conserved, and a short transmembrane motif in the tail (1144-1279 aa) (Fig. 6.1.5).

The detailed search of the translated *Drosophila* genome sequence produced several low score hits. Close analysis of these (searches up and downstream of the respective sequences) revealed they are not true myosins. Despite exhibiting certain similarities to given parts of the myosins none of them showed high homology to a larger part of the head domain. AE003112 showed similarity to the highly conserved region (GESGAGKT) from the P-loop of the head domain. Its polytene chromosome location remains undetermined. AE003614 (CG11199) showed homology to the myosin tail, containing coiled-coil and FERM domains. It is located on chromosome arm 2L (27E), next to the *nina C* gene. AE003495 (CG12047) also showed similarity to the myosin tail, containing coiled-coil and FERM domains. This is located on X chromosome (12E1), close to the newly identified *Myo10A*.



PCR Primers	Expected product	Amplified/Sequenced product
595F-A + 529R	915 bp	915 bp
SECF + 529R	813 bp	813 bp
TRIF + 529R	526 bp	526 bp
595F-A + 595R	1247 bp	1247 bp
SECF + 595R	1145 bp	1145 bp
TRIF + 595R	858 bp	858 bp
595F-B + 595R	496 bp	496 bp
595F-C + LASR	703/2575 bp	764/2636 bp*
PHex + LASR	1076 bp	1076 bp

Figure 6.3.1 RT-PCR analysis of *Mhcl* transcripts

mRNA was prepared from different developmental stages and reverse transcribed. Primers specific for *Mhcl* were used to generate PCR products. The latter were analysed by sequencing and assembled into full length mRNA using the program GeneJockey. The PCR amplification products denoted with * indicate an alternative splicing.

The BLAST search retrieved one more myosin-like gene, *CG15831* in the AE002795 genome scaffolding. This is single exon gene, 219 bp long. The chromosome position of AE002795 has not been previously determined. The program showed similarity to a highly conservative part of the myosin head domain (L**GVLDIFGF**ENFSH**NSFEQ**LCI**NYT**T**NEKLHKFFN**H). We found that the DNA surrounding the gene shows no similarity to the myosin genes (at DNA or protein level). This suggests that AE002795 is either incorrectly assembled in the genome, or the gene *CG15831* is a partial duplication of a myosin gene (this might be either a part of *Mhc* (muscle myosin II) or *didum* (Myosin V) which produced the highest score for this sequence, 59% and 54% identity respectively).

6.4 Discussion

The myosin family has grown significantly in the past decade to encompass more than 177 myosins. These myosins, with a few exceptions fall into 18 classes. Myosins are expressed in both prokaryotes (though these have not been well studied) and eukaryotes.

The annotation of the genome sequences for *S. cerevisiae*, *D. discoideum*, *A. thaliana, C. elegans, D. melanogaster and H. sapiens,* made it possible to identify the complete sets of myosin genes in these organisms. It appears that all eukaryotes have an essential set of three myosin genes, these being from the classes I, II and V, as well as a number of species-specific myosins. In *Drosophila* 13 different myosin genes were identified. The fruitfly has three myosin I genes, two myosin II genes (one encoding for muscle and one for nonmuscle myosins), one myosin III, one V, one VI, two VII, one XV, one XVIII, and one myosin gene that remains to be classified. These classes of myosins have been found in wide range of invertebrate and vertebrate animals. It has been shown that they have a role in a variety of cell functions, including membrane trafficking, signal transduction and maintenance of the cell architecture. Several new myosin genes were identified in the fruitfly. The genome data provided cDNA sequences for four of them (*Myo28B, Myo10A, Mhcl and Myo29D*). The sequences and the open reading frames for these myosins (ORF) were verified by RT-PCR and sequencing. In addition a fifth myosin was found

(*Myo95E*). Using the fruitfly genomic sequence we were able to predict and subsequently isolate and sequence this gene.

Only two of the five newly identified genes in *Drosophila* fell directly into previously known classes, these being Myosin VIIB (*Myo28B*) and Myosin heavy chain like (*Mhcl*). Myosin VIIB is from the well studied class VII myosins with role in the membrane trafficking and stereocilia function.

Mhcl is a member of the recently defined class XVIII of myosins which consists of only two other members, mouse and human PDZ-containing myosins. The PDZ, domain also called DHR (Dlg homologous region), is known to bind either C-terminal or internal (non-C-terminal) polypeptides. These domains have been identified in a broad range of signalling proteins from bacteria, yeast, plants, insects and vertebrates. PDZ domains have been implicated in targeting signalling molecules to sub-membranous sites.

Myosin IC (M*yo95E*) is the third member of class I myosins in *Drosophila*. It contains the Basic Tail domain (TH1) specific to all class I myosins. It is thought that TH1 binds to acidic phospholipids and actin filaments. Interestingly this myosin contains an additional N-terminal insertion, which is similar to the AAA motif, a conserved region of about 220 amino acids that contains an ATP-binding site. This domain is inserted in the region of loop 1 of the motor domain, the other region associated with the hydrolysis site for ATP, and probably modulates its activity.

Myo10A is closely related to class XV myosins. This myosin has a short N-terminal extension which differs from the N-terminal domain characteristic for vertebrate myosins class XV. It also lacks the SH3 and the second MyTH4 domain found in other myosin XV tails. Instead the *Drosophila* Myosin XV tail contains three addition motifs, a GPQ-glycine/proline/glutamine rich domain, a Proline-rich domain and a small Transmembrane domain. The proline-rich sequences have been demonstrated to bind to SH3 (Src homology 3), a small 50 amino acids motif. SH3 domains have been identified in a wide variety of intracellular and membrane-associated proteins and are implicated in signal transduction, linking signals transmitted from the cell surface by protein tyrosine kinases to effector proteins located downstream of the hierarchical pathways.

Myosin 29D is a highly divergent member of the myosin superfamily. Presently it forms a class of its own. This myosin contains an unusual N-terminal extension which shows no homology to other proteins. It also has a small Transmembrane domain in its tail rich in proline residues.

Drosophila melanogaster expresses many myosins genes. Apart from the essential myosins (class I, II and V) it also has myosins from classes III, VI, VII, XV, XVIII, as well as a novel type of myosin. This new data should help to design experiments to investigate the roles of these newly identified myosins in the cell biology and development of *Drosophila*.

Chapter Seven: The function of the *Broad-Complex* during *Drosophila melanogaster* oogenesis

7.1 Introduction

Drosophila oogenesis is an ideal system to study gene regulation during development as it is a comparatively short process with morphologically well defined stages. It takes approximately 3 days for the stage-1 egg chamber to develop into a mature egg. Further, egg chambers, the developmental units of oogenesis, contain only a few types of cells: the somatic follicle cells and the germline cells. The follicle cells perform multiple functions; they undergo dramatic morphogenetic movements and interact with the germline cells to establish both the anterior/posterior (A/P) and dorsal/ventral (D/V) axes of the egg and embryo, as well as producing yolk proteins for the developing oocyte and then chorion (the protective shell that enables the egg to survive in the environment).

Pattern formation and cell fate specification during development require complex regulation of the multiple families of transcription factors to establish and coordinate transcriptional cascades. Transcription factors are expressed differentially in a temporally ordered manner governed by a genetically based hierarchical time table. They in turn activate or repress the tissue specific expression of structural genes, a critical process in morphogenesis. One way to regulate the temporal and spatial expression of developmental genes is by alternative splicing. Thus different messenger RNAs can be generated from identical pre-messenger RNA transcripts. This is the primary mechanism by which the *Broad complex*(*BR-C*), an early ecdysone response gene, is regulated.

The *Broad-Complex* encodes a family of zinc-finger transcription factors (DiBello et al. 1991; Bayer et al. 1996a) and was shown to be expressed in the follicle cells in a dynamic pattern. The late pattern (stage 10B of oogenesis) is defined by two groups of dorsal-anterior follicle cells (Deng and Bownes 1997) and is specified by the epidermal growth factor receptor (EGFR) signalling pathway and further regulated by decapentaplegic (DPP) signalling pathway along the D/V axis. This localised expression is associated with the function of the *BR-C* in dorsal appendage formation. The involvement of *BR-C* in dorsal appendage morphogenesis was shown by mutational analysis of *BR-C* partial "loss-of-function" mutants, and

was supported by ectopic expression of *BR-C* "transgenes" during oogenesis. It is thought that the *BR-C* may provide a link between pattern formation and cell differentiation in morphogenesis (Deng and Bownes 1997).

7.1.1 Molecular organisation of Broad complex gene

The *BR-C* encodes a family of C2H2 zinc-finger proteins (Z1, Z2, Z3, and Z4), which share a common amino terminal (the *BR-C* "core") domain, but differ in zinc-finger domains (DiBello et al. 1991). The core is alternatively spliced to one of four pairs of zinc-finger encoding exons (Figure 7.1.1A, B), generating four classes of transcripts. Each transcript class is translated into a specific protein izoform: Z1, Z2, Z3 or Z4. In addition, three Z1 isoforms have been identified: Q^1-Q^2-Z1 encodes 232 amino acids which are fused to the BR-C core domain (Q stands for glutamine rich region); Q^2-Z1 lacks the first 17 amino acids from the glutamine-rich part of the Z1 exon (this is due to alternative splicing, utilising a shorter version of the Z1 exon, while preserving the same open reading frame); TNT-Q^1-Q^2-Z1 contains a small threonine, asparagine rich (TNT) exon which is preceding the Z1 exon but is not part of the BR-C core domain (DiBello et al. 1991; Tzolovsky et al. 1999).

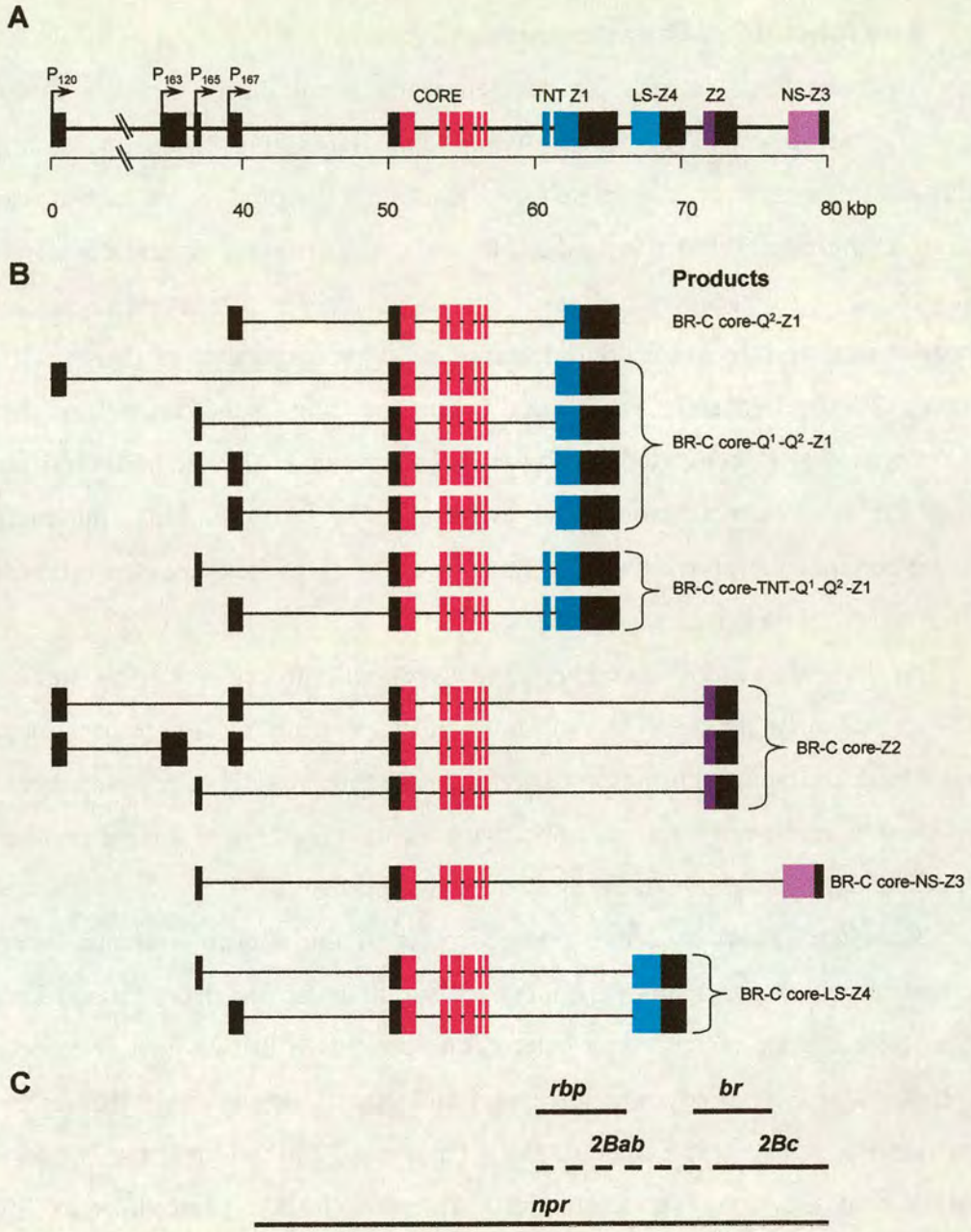


Figure 7.1.1 Organisation of the *BR-C* gene

A: Molecular organisation of the *BR-C* gene, which maps to the 2B5 region (adapted from (Chao and Guild 1986) and (Gonzy et al. 2002). For simplicity and clarity the first intron of the *BR-C*, is not shown on the figure. Coloured boxes represent open reading frames; black filled boxes represent untranslated regions of the *BR-C* transcripts. Four putative promoters have been previously described: P Distal at nucleotide 120; and P Proximal at nucleotides 163, 165 and 167 (DiBello et al. 1991; Bayer et al. 1996a). Z1 and Z2 transcripts can be synthesised from either P_{120}, P_{165} or P_{167}, while Z3 and Z4 only initiate at P_{165} and P_{167} (Bayer et al. 1996a). Z2 is the only transcript that uses P_{163} start site.

B: Organisation of zinc-finger isoforms. Differentially spliced *BR-C* transcripts share a common Core domain linked to one of four (Z1-Z4) pairs of C2H2 zinc-finger domains.

C: Complementation map of the *BR-C* based on (Bayer et al. 1997; Gonzy et al. 2002).

7.1.2 The function of Broad complex

Metamorphosis is a major developmental event during the life cycle of holometabolous insects. The relatively immobile larva transforms into a highly motile adult with systems specialised for dispersal and reproduction. Larval tissues undergo a controlled hystolysis as the imaginal disks differentiate and develop into adult structures. This sequence of tissue specific changes is driven and regulated by fluctuations in the titre of the steroid hormone 20-hydroxyecdysone (Buszczak and Segraves 2000). Ecdysone is released from the ring gland in pulses during development, and is converted to 20-hydroxyecdysone in the fat body and target tissues. Prior to each developmental event such as larval molting, pupariation, pupation and metamorphosis the concentration of the 20-hydroxyecdysone (hereafter referred as ecdysone) rises.

At the end of larval development in *Drosophila* an ecdysone peak induces a small set of early polytene chromosome puffs. Puffing results from structural changes in the polytene chromosomes and indicate gene activity. The products of the genes located in the early puffs are necessary for the induction of a large number of late puffs and the repression of the early ones (Ashburner 1974).

Ecdysteroids act by binding to receptors of the nuclear hormone receptor superfamily, which are ligand-dependent transcription factors (Henrich and Brown 1995). The ecdysone receptor is a heterodimer composed of two nuclear receptors, the products of *EcR* (Ecdysone Receptor) and *Usp* (Ultraspiracle) (Horner et al. 1995; Henrich et al. 1994; Yao et al. 1992; Grad et al. 2001). Usp is the *Drosophila* homologue of the mammalian retinoid X receptor (RXR) (Ghbeish et al. 2001; Zelhof et al. 1997). The EcR/Usp dimer can bind ecdysone-responsive regulatory DNA sequences, thus repressing transcription in the absence of hormone and vice versa. There is evidence that Usp is also able to dimerise with other steroid receptors (Sutherland et al. 1995; Zelhof et al. 1995).

The *Broad complex (BR-C)* has been identified as an essential gene required for *Drosophila* metamorphosis (Karim et al. 1993). *BR-C* is located within the limits of the early ecdysone inducible puff 2B4-5 on the X chromosome. It is one of three early ecdysone responsive genes, including *E74* and *E75* (Burtis et al. 1990;

Segraves and Hogness 1990). These genes encode transcription factors, which are directly activated by the ecdysone receptor and co-ordinate the subsequent transcription of secondary effector genes (Ashburner 1974); and for reviews see (Bayer et al. 1996b; Kiss et al. 1988; Zhimulev et al. 1995; Karim et al. 1993). All three early genes are unusually long: *BR-C* ~ 80 kbp; *E74* ~ 17 kbp, *E75* ~ 108 kbp, and generate a large number of spatially and temporally regulated alternative transcripts.

Genetically, the *BR-C* locus has three complementation groups, defined by mutations: *br* (shorter and wider wings), *rbp* (reduced bristle number on palpus), and *2Bc* (causes early pupal developmental arrest). An additional *2Bab* group does not complement either the br and rbp mutations but does complement the *2Bc* mutations. A group of mutations that do not complement any of the other mutations defines a non-pupariating (*npr*) class (Figure 7.1.1C) (Belyaeva et al. 1980; Kiss et al. 1988). Nonpuparating mutants are phenotypically indistinguishable from deletions of the *BR-C* locus (they correspond to a complete loss of function) and die as wandering third instar larvae, failing to puparate. Genetic analysis of *BR-C* has shown that the gene is essential for the morphogenesis of imaginal discs. The function of br^+ is primarily required for the elongation and eversion of the imaginal discs into adult appendages as well as tanning and hardening of the larval cuticle. The rbp^+ function, has been implicated in muscle and bristle development. In addition, both rbp^+ and $2Bc^+$ functions are essential for histolysis of the larval tissues and gut morphogenesis. It was found that $2Bc^+$ has important role in the fusion of the imaginal discs to form a continuous adult epidermis, (Kiss et al. 1988). All three functions are also required for the reorganisation of the CNS (Kiss et al. 1988; Emery et al. 1994).

It was suggested there may be a one-to-one link between the genetic function of a given complementation group and a specific protein isoform (Z1-Z4) (DiBello et al. 1991). Current data shows that the relationships between the complementation groups and protein isoforms are more complicated. It several case it has been demonstrated that one isoform can provide a regulatory function opposite that of the other isoform (Hodgetts et al. 1995; von Kalm et al. 1994). Clearer data on these relationship were provided by Bayer et al. (1996a), who showed that the lethality

associated with each of the complementing groups was rescued using heat-inducible transgenes. This data was supported by (Gonzy et al. 2002) with the isolation of a number of novel mutations in the *BR-C*. It was found that br^+ function is only provided by the Z2 isoform. Despite this, there may be functional redundancy or regulatory dependency associated with rbp^+ and $2Bc^+$ functions (Bayer et al. 1996a). The Z1 isoform was shown to provide full rbp^+ function, while Z4 provides partial function. So far no mutants for Z4 have been identified and its function remains unclear. The *2Bc* lethality is fully rescued by heatshock expression of a Z3 transgene, and partially rescued by a Z2 transgene.

7.1.3 Functional domains in Broad Complex

7.1.3.1 BTB domain

The BR-C core contains a highly conserved amino-terminal motif of approximately 115 amino acids called the BTB domain. In *Drosophila* there are at least 40 members of the BTB family which appear to be involved in protein-protein interactions (homodimerisation and heterodimerisation) and are widely distributed among metazoans (Bardwell and Treisman 1994; Zollman et al. 1994; Couderc et al. 2002). The family was named after the first three genes in *Drosophila* found to contain BTB domain: *BR-C* , *tramtrack* and *bric a brac* (Figure 7.1.3A). The domain is also known as POZ (poxvirus and zinc finger).

The Ttk protein (*tramtrack*) binds to the promoter of the pair-rule genes *fushi tarazu* and *even-skipped* (Harrison and Travers 1990; Brown and Wu 1993) and is responsible for the repression of these genes during embryogenesis. The gene is also expressed downstream of *Notch* with a function in the specification of daughter cell fates during the assymetric cell divisions in the peripheral nervous system (Guo et al. 1995). Lai and Li (1999) have demonstrated a positive role of one of the Ttk isoforms (Ttk69) in the differentiation of the photoreceptor neurons at the later stages of *Drosophila* eye development.

The gene *bric à brac* (BAB) is essential for pattern formation along the proximal-distal axis of the leg and antenna (Zollman et al. 1994). The gene was also found to have important function in the basal and interfollicular stalk development of the ovariole (Godt and Laski 1995). Null mutations in *bab* disrupt ovarian morphogenesis and lead to female sterility.

Figure 7.1.3 Protein alignment of the BR-C domains

The alignments were performed with CLUSTAL W (Thompson et al. 1994) available from the WEB-based package at http://www.ebi.ac.uk/clustalw/. Subsequently, the multiple alignments were analysed and shaded on the BoxShade server (http://www.ch.embnet.org/ software/BOX_form.html). The sequences were exported to Microsoft Word as RTF-new files. The consensus sequence is given in red.

A: BTB domain alignment. The sequence of the BR-C BTB domain is compared to BTB domains from Ttk (*tramtrack*) and Bab (*bric a brac*); NCBI accession numbers: BAB-I (CAB64385), BAB-II (CAB64388), TTK (P42282), and BR-C (Q01295)

B: Protein alignment of BR-C zing finger domains Z1 (X54666), Z2 (X54665), Z3 (X54664), Z4 (U51585), AZ2 (NP_610289), Ttk69 (NP_524911), and Hunchback (O62537).

7.1.3.2 Zinc finger domains

Zinc finger motifs are small DNA binding domains. They exist as different types, depending on the positions of the cysteine residues. In BR-C the zinc fingers are from the C2H2 type, composed of ~30 amino acids including two conserved Cystidine and two conserved Histidine residues that fold in a complex tertiary structure with a central Zn^{2+} ion. This arrangement yields a finger-like projection that protrudes from the zinc coordination site and is responsible for the nucleic acid binding properties of the domain. It was shown that C2H2 zinc fingers recognise

short, approximately 5 base pairs, G-rich nucleic acid sequences (DNA or RNA). Figure7.1.3B shows a protein alignment of the BR-C zinc finger domains compared to C2H2 fingers from several other *Drosophila* genes known to contain zing finger domains.

7.1.4 Chorion genes in *Drosophila*

The somatic follicle cells synthesise endochorion and exochorion, the two outer layers of the *Drosophila* eggshell, during the last 6 hours of the oogenesis. Five chorion genes are clustered on two genetic loci, one on the X-chromosome (polytene position 7E9) encoding Cp36 and Cp38, and one on the Third chromosome (polytene position 66D14), producing Cp15, Cp16, Cp18 and Cp19 (Waring and Mahowald 1979; Spradling and Mahowald 1980). The chorion proteins are transported and secreted at the inner surface of the follicle cells from where they are incorporated into the eggshell.

During late oogenesis the two clusters of chorion genes are selectively amplified, above the level of the remainder of the follicle cell genome, which also endoreplicates to produce polyploid cells (Spradling and Mahowald 1980; Spradling et al. 1980; Orr-Weaver 1991). In response to developmental signals in the follicle cells the third chromosome cluster is amplified 60-80 fold and the X-chromosome set, 15-20 fold above the rest of the genome. This specific amplification depends on cis-acting sequences amongst the chorion genes (Orr-Weaver and Spradling 1986; Delidakis and Kafatos 1989).

7.2 Project aims

Over the last several years a number of developmentally important questions about cell activity and function of the somatic follicle cells have been studied. The data suggests that a crucial role is played by the follicle cells in oocyte determination, maturation and in polarity establishment. However, there are many more questions that remain unanswered: how is the oocyte determined, what signals trigger follicle cell differentiation and what are the interactions between the subtypes of follicle cells and between the follicle cells and the germ line cells?

The aim of this work was to study the early function of *Br-C*, when it is expressed in all follicle cells at stage 6 of oogenesis. Its later expression in the dorsal-anterior follicle cells, is known to be related to its function in dorsal appendages morphogenesis (Deng and Bownes 1997). Since the chorion genes encode major eggshell components, and *rbp+* function has been reported to be necessary for chorion gene amplification during oogenesis (Orr et al. 1989; Huang and Orr 1992), the purpose of this work is to investigate the relationship between the *Br-C* and chorion gene amplification and expression.

7.3 Project background

7.3.1 Isolation and expression of the *BR-C* gene in oogenesis

An extensive P[Gal-4] enhancer trap screen for genes involved in D*rosophila* oogenesis was carried out in our laboratory (Deng et al. 1997). *Br-C* was identified as the target gene in one of the lines. The reporter gene showed specific expression patterns in two sets of dorsal-anterior follicle cells at stage 10 of oogenesis. This finding prompted further investigations.

In situ experiments revealed that Z1 is the sole zinc-finger isoform expressed at detectable levels during oogenesis. The *Br-C* is initially expressed in all follicle cells at stage 6 (Figure 7.7.1A). Later at stage 10, *Br-C* forms a gradient of expression with the strongest signal localised to the dorsal-anterior follicle cells (Figure 7.7.1B). Shortly after that (stage 10b, 11) the *Br-C* transcripts are detected in just two sets of dorsal-anterior follicle cells (Figure 7.7.1C), which are thought to be the precursors of the chorionic appendages (Deng and Bownes 1997).

Immunochemical analysis with antibodies for the Z1 isoform showed that the BR-C protein distribution pattern is similar to that of Z1 mRNA during stages 6-8 of oogenesis, when all follicle cells stain. The BR-C (Z1 isoform) is then detected at stage 10 with staining in all columnar follicle cells except the dorsal anterior cells, a pattern similar to the mRNA distribution (Deng and Bownes 1997). However, the follicle cells at the posterior pole appear to be stained at this stage, which does not overlap completely with the mRNA distribution pattern. The late distribution pattern of the *BR-C* protein and mRNA differs. At stages 11-12 a strong signal is observed in

two groups of lateral-dorsal follicle cells but the posterior and ventral follicle cells are still stained. The signal in the posterior and ventral region gradually decreases and disappears at late stage 13, leaving only two patches of dorsal-anterior follicle cells stained. A likely reason for that difference is that the half-life of the protein is much longer than that of the mRNA. Thus, by the time when late *BR-C* transcription occurs in the lateral-dorsal-anterior follicle cells, the protein translated from the early *BR-C* transcripts are still present in the posterior and ventral follicle cells, while the mRNA has been degraded. This explains how the early and late protein distribution patterns overlap to form a gradient pattern at stages 11-12. This can also explain why the protein, but not the mRNA, is detected in the follicle cells at the posterior pole during stage 10. Antibodies that recognise the BR-C core domain produced similar a staining pattern to Z1 antibodies, while Z3 antibodies did not generate any detectable staining during the oogenesis. Antibodies to Z2 and Z4 were not available at that time and their expression patterns have not been investigated.

Deng and Bownes (1997) also showed that BR-C is expressed only in the nuclei of the follicle cells, which is consistent with the fact that the *BR-C* encodes a transcription factor.

7.3.2 The function of *Broad Complex* during *Drosophila* oogenesis

Genetical analysis of partial "loss of function" mutants indicated that *Br-C* is required for dorsal appendage morphogenesis during oogenesis. Deng and Bownes (1997) showed that signals from both the germline and the somatic follicle cells are required to determine the correct *BR-C* expression pattern. First, the Grk (TGF-α homologue) signal from the oocyte binds to the epidermal growth factor receptor (EGFR) to activate a tyrosine kinase signalling pathway which establishes a dynamic pattern along the D/V (dorsal/ventral) axis in a dose dependant manner. The patterning along the A/P (anterior/posterior) axis is directed by TGF-β/DPP (Decapentaplegic) signalling pathway. These two pathways (EGFR and DPP) cooperate to specify the expression of *Br-C* into two discrete groups of lateral-dorsal-anterior follicle cells. These then differentiate and direct the formation of the dorsal appendages in the correct position on the eggshell (Deng and Bownes 1997).

A number of experiments were done by Wu Min Deng to determine which BR-C isoform is required for dorsal appendage formation (Deng, Thesis, 1997). The Z1 zinc-finger is the only isoform expressed at levels detectable by *in situ* hybridisation and immunohystochemistry during oogenesis. Bayer et al. (1997) reported that Z1 provides the full rbp^{+} function. To test whether the rbp^{+} (Z1 isoform) function is responsible for the dorsal appendage formation, female homozygous viable rbp^{1} and rbp^{2} mutants (escapers) (Table 7.4.1) were dissected and the eggshell phenotype examined. It was found that the dorsal appendages are shorter and irregular in shape compared to wild type suggesting that rbp^{+} is indeed involved in the dorsal appendage morphogenesis.

7.4 Materials and fly strains used in the analysis of the Broad complex

7.4.1 List of the *Drosophila* strains used to study *Broad complex*

Oregon R and W^{1118} (the host strain used in the construction of the transgenic heatshock lines) were used as wild-type controls. The rbp^2 and Br^1 alleles are viable and were maintained as homozygotes. All other *BR-C* mutations were maintained over *Binsn, an* X-chromosome balancer carrying the markers *Bar* and *singed*. Since all mutant and transgenic stocks showed slightly reduced viability they were maintained on standard corn meal food at 26°C

Table 7.4.1 *Drosophila* strains

Stock no	Genotype	Comments	Reference
br^1	*br^1*	deficiency line	(Kiss et al. 1988)
br^5	*$l(1)t^{35}$/ Binsn*	deficiency line	(Kiss et al. 1988)
br^6	*$l(1)t^{103}$/ Binsn*	deficiency line	(Kiss et al. 1988)
rbp^1	*br^{rbp1}/ Binsn*	deficiency line	(Kiss et al. 1988)
rbp^2	*br^{rbp2}/ Binsn*	deficiency line	(Kiss et al. 1988)
$2Bc^1$	*br^{Bc1}/ Binsn*	deficiency line	(Kiss et al. 1988)
$2Bc^2$	*br^{Bc2}/ Binsn*	deficiency line	(Kiss et al. 1988)
npr^6	*br^{npr-6}/ Binsn*	deficiency line	(Kiss et al. 1988)
br^{A47}	*br^{A47}/ Binsn*	deficiency line	(Deng and Bownes 1997)
hsZ1 (527)	*w; P{BRcore-Q1-Z1}*	carries Q^1-Q^2-Z1 transgene	(Bayer et al. 1997)
hsZ1 (708)	*w; P{hs-BRC-Z1}*	carries TNT-Q^1-Q^2-Z1 transgene	(Bayer et al. 1997)
hsZ2 (CD5)	*w; P{hs-BRC-Z2}*	carries Z2 transgene	(Bayer et al. 1997)
hsZ3 (797)	*w; P{hs-BRC-Z3}*	carries Z3 transgene	(Bayer et al. 1997)
hsZ4 (28-I)	*w; P{hs-BRC-Z4}*	carries Z4 transgene	(Bayer et al. 1997)

7.4.2 PCR analysis of *Broad complex* transcripts

BR-C transcripts levels in ovaries were detected by RT-PCR as described previously (Hodgetts et al. 1995). Total RNA from ovaries and larvae (control RNA) was isolated using RNAeasy-Total RNA Kit (see methods, 2.5.12). The RNA (5 μg) was primed with Oligo-p$(dT)_{15}$ and reverse transcribed using Superscript^{TM} II (methods, 2.5.10). For the subsequent DNA amplification, 5% of the first strand reaction mix was used. To amplify each of the zinc-finger domains, appropriate primer pairs (Table 7.4.2) were added to the PCR mixture: a common primer for the core domain was combined with one of the four primers for the respective zinc-finger motif. The sequence data for the primers was obtained from (DiBello et al. 1991; Hodgetts et al. 1995) (Table 7.4.2). In the PCR reaction QIAGEN Taq Polymerase and the protocol designed for use with Q-Solution was used. The PCR reaction was carried out as follows: one cycle at 94°C for 4 minutes; 35 cycles, step one - at 94 °C for 30 seconds, step two - 60 °C for 30 seconds, step three - 72°C for 1.5 minutes; one cycle at 72°C for 7 minutes.

Table 7.4.2 Broad complex PCR primers

Primer name	Sequence in 5' > 3' direction	Length (nt)	Tm (°C)
Z1	TGC TGG TGC TGC TGG TGA TA	20	62
Z2	TCA TCT CCA TTT CGC CGG GA	20	62
Z3	GAT GGC GGT CGT CTT AAG CA	20	62
Z4	GTG GTT GTT CAG CGA GTT CA	20	60
CORE 3'	ACA AGA TGT TCC ATG CAG CC	20	60

7.5 *rbp*$^{+}$ function is required for dorsal appendage formation

To further test if *rbp* is the major functional domain involved in dorsal appendage formation we did a genetic analysis using the *npr* allele in crosses with *br*, *rbp*, and *2Bc* alleles (Table 7.4.1). The cross between rbp^{1} and npr^{6} produced only two males and no viable female heteroallelic mutants (Table 7.5). Two rbp^{2}/npr^{6} females were produced in the cross with rbp^{2} flies. They lived for two days without laying any eggs. Then they were dissected to examine the ovarian phenotype. It was found that the ovaries were not completely developed and the few late stage oocytes formed had no appendages (Wu Min Deng was also involved in these experiments). Thus we were unable to examine *rbp* function by this method. The combination $2Bc/npr^{6}$ was found to be completely lethal, and as a result we could not establish if there is a function encoded by Z3 (2Bc) isoform (Table7.5).

The function of Z2 was investigated further. The cross between br^{5} and npr^{6} produced no viable heteroallelic flies, while the cross between br^{6} and npr^{6} generated 22 heteroallelic males, but no females which could be examined. It was observed that eggs produced by br^{1}/br^{A47} and br^{1}/npr^{6} mothers have reduced dorsal appendages (Deng and Bownes 1997), suggesting that the *br* functional domain is likely to be required for dorsal appendage formation. In order to test this hypothesis, eggs produced by br^{5}/br^{1} and br^{6}/br^{1} females were examined, and found to have normal dorsal appendages. This observation, along with the fact that the br^{A47} is actually an *rbp* allele (Huang and Orr 1992), suggests that the *br* functional domain, and hence Z2, is not involved in dorsal appendage formation.

Table 7.5 Survival frequencies of *BR-C* mutants

Cross	Progeny			
	Heterozygotes[a] +Homozygous Balancer[b]	Heteroallelic mutants[c]		Lethal period
		♂	♀	
$\frac{br^1}{br^1} \times \frac{npr^6}{Binsn}$	216	21	9	not observed
$\frac{br^5}{Binsn} \times \frac{npr^6}{Binsn}$	158	0	0	III instar larva
$\frac{br^6}{Binsn} \times \frac{npr^6}{Binsn}$	36	22	0	pupal
$\frac{rbp^1}{Binsn} \times \frac{npr^6}{Binsn}$	140	2	0	pupal
$\frac{rbp^2}{rbp^2} \times \frac{npr^6}{Binsn}$	130	11	2	pupal
$\frac{2Bc^1}{Binsn} \times \frac{npr^6}{Binsn}$	73	0	0	II-III instar larva
$\frac{2Bc^2}{Binsn} \times \frac{npr^6}{Binsn}$	215	0	0	pupal

[a]Heterozygotes, flies carrying a *BR-C* mutant chromosome and a balancer chromosome (*br/Binsn, rbp/Binsn, 2BC/Binsn*, and *npr/Binsn*).
[b]Homozygous balancer, flies homozygous with respect to the balancer chromosome (*Binsn/Binsn*).
[c]Heteroallelic mutants, flies carrying different combinations of *BR-C* alleles (*br/npr, rbp/npr, and 2BC/npr*).

How can the phenotype of br^1/br^{A47} and br^1/npr^6 eggshells be explained if *br* is not the functional domain required for dorsal appendage formation? This could be understood if the br^1 mutant not only affected *br* function, but also affected *rbp* function. In order to test this possibility, the eggshell phenotype of eggs laid from br^1/rbp^1 mutants was examined. It was shown that eggs produced by the br^1/rbp^1 mothers have reduced dorsal appendages, similar to those produced by the br^1/br^{A47} females. This indicates, therefore, that the br^1 is in fact a weak *2Bc* or *2Bab* allele, which fails to complement either *rbp* or *br* function. This suggests that *rbp* (Z1 and Z4 isoforms) is a functional domain involved in dorsal appendage formation during oogenesis, however, we cannot rule out the involvement of Z3 from these experiments due to the failure of these crosses to generate adult females due to early lethality.

7.6 *Ectopic BR-C expression induces ectopic dorsal appendage material*

Is *BR-C* function sufficient to direct the formation of the dorsal appendages? To address this, heat-inducible Z1 transgenic flies (*hsp70*/Z1) were used to examine the effect of ectopic *BR-C* expression during oogenesis (Table 7.4.1). Following standard heatshock (37°C, 30min) and incubation at 26°C for 2 to 48 hours, ovaries of the *hsp70*/Z1 females were analysed (Table 7.6). Flies were dissected to examine the effect of the heatshock at 2, 3, 5, 9, 24, and 48 hours after heat treatment. The first abnormalities in the egg chambers were observed 3 hours after the heatshock. The eggs laid during the first 3 hours following heatshock also have a very high hatch rate, presumably being sufficiently differentiated at the time of the heatshock for the ectopic *BR-C* expression to have any effect. The results over this period do not differ significantly for the control heatshocks (Table 7.6). The strongest effect was observed between 4 and 6 hours. It was observed that extra dorsal appendage material was produced in the dorsal-anterior region of the eggshells (Figure 7.6B-F). In most cases, dorsal appendage material appeared in the dorsal gap between the two appendages. It condensed at the base of the dorsal appendages and less material was deposited in the appendages themselves. The dorsal appendages did not elongate properly (Figure 7.6B-D) presumably due to a failed migration of the follicle cells. Different phenotypes have been observed depending on the stage of the egg chamber at the time of heatshock. Heatshock at the time of dorsal appendage formation, stage 11, leads to an "appendageless" phenotype or to small fused appendages (Figure 7.6B, C). Appendages with abnormal shapes and/or different lengths have been observed on eggs staged 12-13 at the time of heatshock (Figure 7.6D, E, F). These observations indicate that ectopic Z1 can induce formation of ectopic dorsal appendage material. Nevertheless, the ectopic dorsal appendage material is restricted to the dorsal anterior eggshell, suggesting that the fate of the follicle cells is pre-determined along the two major axes prior to the requirement for *BR-C* function in this process. It was hypothesised that the lack of *BR-C* expression in the dorsal-most follicle cells is due to high levels of expression of *pointed (pnt)* in those cells (Deng and Bownes 1997). The data shown here indirectly support this hypothesis. In the dorsal-most follicle cells, there could be competition between the expression of Pnt

and *BR-C*. When high levels of Pnt are expressed, *BR-C* expression is inhibited in these cells. However, in heat-shock lines, the *BR-C* expression would overcome the inhibition by Pnt. Thus, dorsal appendage material can be synthesised by these cells.

Table 7.6 Effect of heatshock on Z1-Z4 transgenic flies

Experiment	Fly stock	No of eggs laid	Percentage of abnormal eggs	Percentage of eggs hatched
0-3h	Z1	172	5.8	94
	Z2	144	2.3	96
	Z3	134	1.5	98
	Z4	80	5.0	96
	w^{1118}	364	1.6	98
	OrR	94	2.1	99
3-5h	Z1	27	74	2
	Z2	68	55	8
	Z3	4	0	25
	Z4	23	43	4
	w^{1118}	137	2.5	97
	OrR	117	2.7	96
5-24h	Z1	10	89	6
	Z2	361	78	12
	Z3	27	42	18
	Z4	186	77	15
	w^{1118}	505	4.4	94
	OrR	363	3.6	98
24-48h	Z1	9	62	72
	Z2	670	33	93
	Z3	72	24	90
	Z4	390	18	97
	w^{1118}	520	1.3	100
	OrR	370	1.1	99

Note: Z1-Z4, transgenic flies carrying constructs to misexpress the various zinc-finger isoforms of the *BR-C* in response to heatshock; W^{1118} is the host line used for the construction of the transgenic lines used as a control. OrR (wild-type) flies were also used as a control. Approximately 100 flies (50 males and 50 females) from each stock were used in the heatshock experiments.

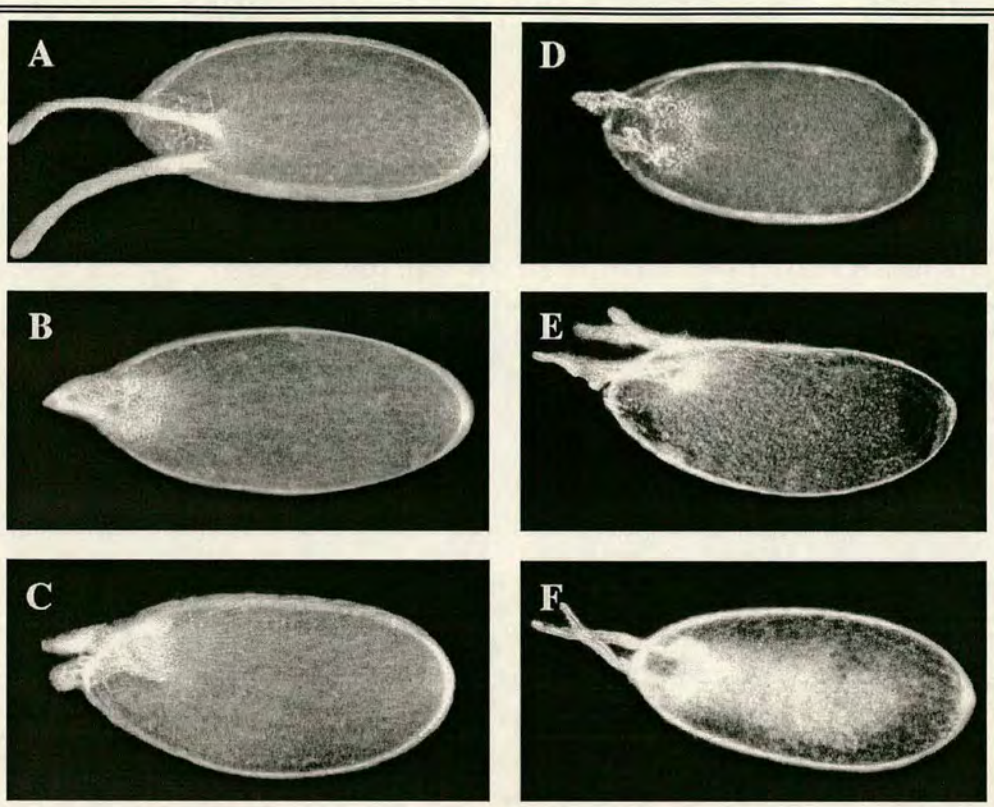


Figure 7.6 Ectopic expression of different *BR-C* zinc-finger isoforms during oogenesis

A-F: dark field microscopy;

A: Shows a wild type egg after heatshock treatment.

B,C: Two distinctive *hsp70/Z1* (527-5 and 708-1) transgenic lines exhibit a similar phenotype. Dorsal appendage materials are present in the dorsal-most region, resulting in fused, thickened and shortened dorsal appendages.

D-F: Shows a similar, but less severe eggshell phenotype in *hsp70*/Z1 flies-see explanation in the text. Fused, thickened, and irregularly shaped appendages are also found in eggshells produced by the Z2, Z3 and Z4 transgenic lines.

Although Z1 seemed to be the sole *BR-C* zinc-finger isoform expressed at high levels during oogenesis when analysed by *in situ* hybridisation, we tested Z2, Z3 and Z4 to determine if they exhibit a similar phenotype when ectopically expressed during oogenesis. Thus, *hsp70*/Z2, *hsp70*/Z3 and *hsp70*/Z4 flies were heat-shocked and the eggshell phenotype was examined (Table 7.6). It was found that ectopic dorsal appendage material is produced in the dorsal-anterior region of the eggshells by all three transgenic lines. This phenotype is similar to that exhibited by

eggs by the *hsp70/Z1* flies after heatshock, suggesting that all of the four zinc-finger isoforms could be functional in dorsal appendage formation during oogenesis.

It is apparent from Table 7.6 that heatshock has the strongest effect on chorion morphology and egg viability in *hsp70*/Z1 flies. Z2-Z4 recovered viability to approximately 95% in two days, while Z1 recovered only to 72% during that time period. It was also found that heatshocked Z1, Z2 and Z4 flies lay abnormal eggs (Table 7.6). The ectopic expression of *BR-C* in Z3 flies was found to disrupt the process of egg development soon after the heatshock. We observed that some 20% of all laid eggs have aberrant micropyles, due to excess chorion formation. This could prevent the sperm entering the egg and hence subsequent development would fail due to lack of fertilisation. Another possible explanation is that ectopic expression of *BR-C* can disrupt some other BTB containing protein that can dimerise with the *BR-C* and thus modulate its function.

Heatshock alone causes eggshell defects. The data of the control experiments with the heatshocked wild type OrR flies and w^{1118} the host line for the transgenic flies is presented in Table 7.6. We observed in the few abnormal eggs wide branched dorsal appendages of approximately normal length. It is quite clear that the results of misexpressing *BR-C* in oogenesis significantly affects the eggshell.

7.7 Other zinc finger isoforms are expressed in oogenesis

Although only Z1 expression was clearly detectable by *in situ* hybridisation we observed defects in chorion formation and morphology by overexpressing all four zinc finger isoforms available. It became essential, therefore, to establish if this was due to some degree of functional redundancy between the isoforms with respect to eggshell development or if the other zinc fingers are, in fact, expressed at lower levels in oogenesis.

Firstly we repeated the *in situ* hybridisations to whole mount ovaries, following a modified protocol, using ssDNA probes labelled by PCR (see methods, 2.5.14). We observed that Z1 transcripts are expressed in a dynamic pattern (Figure 7.7.1 A, B, and C) consistent with the previously reported data (Deng and Bownes 1997). No transcripts for Z2 and Z3 isoform were detected by *in situ* hybridisation

(Figure 7.7.1D and E). We observed that Z4 is initially expressed in all follicle cells at late stage 5 to stage 6. Later at stage 10, low levels of Z4 transcripts were detected in the columnar follicle cells (Figure 7.7.1F). This showed that the expression pattern of Z4 is similar the expression pattern of Z1. However the transcripts for Z4 isoform did not localise specifically to the lateral-dorsal-anterior follicle cells to produce the characteristic pattern for the Z1 isoform.

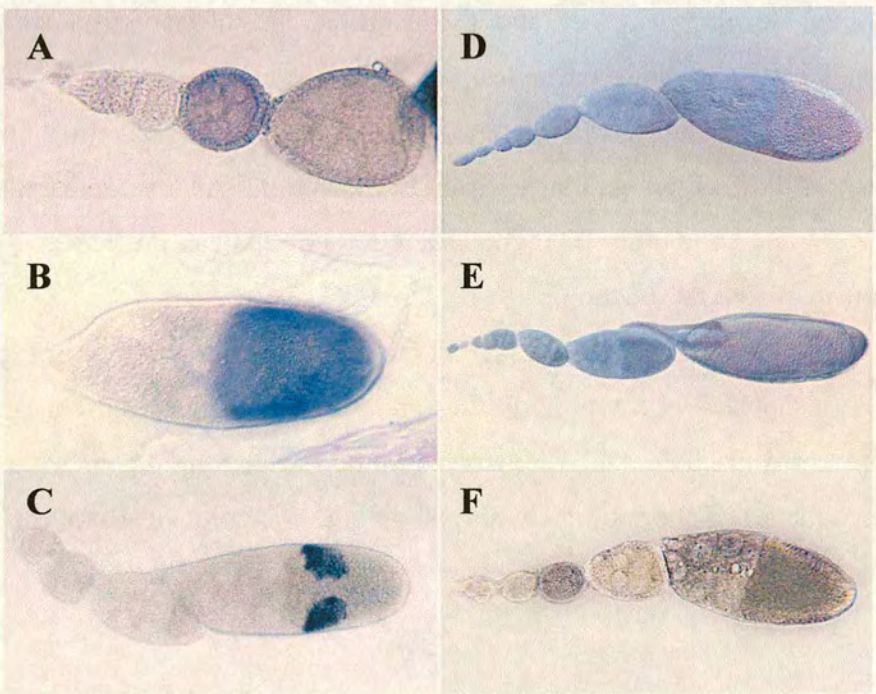


Figure 7.7.1 *BR-C* expression pattern in wild-type ovaries
A: *BR-C* transcripts (Z1 isoform) are first detected in all follicle cells, at stage 6 (black arrows).
B: At stage 10A most of the columnar follicle cells express *BR-C* (Z1), except the dorsal anterior region and the cells at the posterior pole.
C: At stage 10B The *BR-C* (Z1) staining is confined to two discrete patches of follicle cells lying symmetrically on either side of the dorsal midline. The distance between the stained patches is approximately 3 cell wide (blue arrow).
D, E: No expression was detected using probes for Z2 and Z3 isoforms.
F: *In situ* hybridisation with probe for Z4 transcripts produced a low level staining similar to the Z1expression pattern.

RT-PCR was performed to detect low levels of *BR-C* transcriptional activity (for details see 7.4.2). To check the expression of the different transcripts we used primer pairs for the different isoforms and the common core domain (the organisation of the zinc finger isoforms in relation to the *BR-C* gene are shown in Figure 7.1.1). The position of the PCR primers on the *BR-C* map is illustrated on Figure 7.7.2, while their sequence is given in Table 7.4.2.

The results clearly show that all 4 zinc fingers are expressed in oogenesis (Figure 7.7.3). The identity of the PCR generated products was confirmed with Southern blot analysis (Figure 7.7.4).

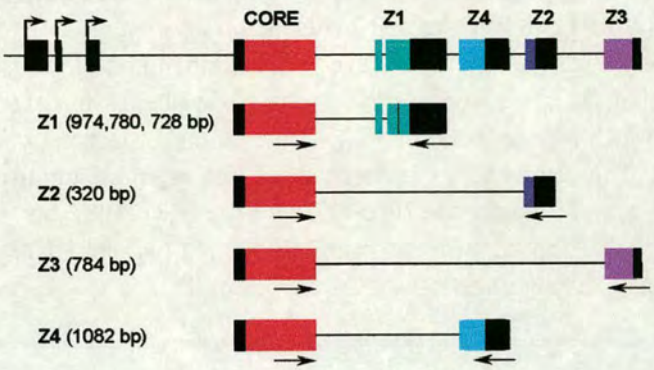


Figure 7.7.2 Exon map of *BR-C*

Simplified exon map of *BR-C*. The upstream half of gene, including the first exon is omitted from the figure. Arrows indicate the relative position of the primer pairs used in the RT-PCR analysis. All four PCR reactions were carried out with a common primer located at the 3′ end of the core domain of *BR-C*. The other primers are located just within, or immediately 3′ to the respective Z1-Z4 domain. The predicted sizes of the PCR amplification products are given beneath the map. The primers used should generate products of 974 bp, 780 bp, 728 bp (Z1); 320 bp (Z2); 784 bp (Z3); and 1082 bp (Z4) respectively, based on published data (Hodgetts et al. 1995; Bayer et al. 1996a) (Figure 7.7.3).

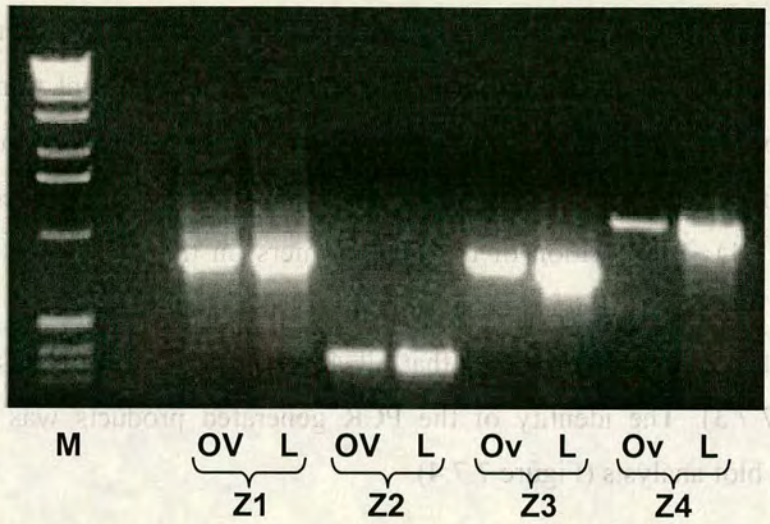


Figure 7.7.3 PCR analysis of the *BR-C* transcripts in *Drosophila melanogaster* (wild-type, Oregon R) ovaries.

Total RNA from ovaries was reverse transcribed and the subsequent cDNA PCR amplified using the primer pairs shown with arrows below the exon map (Figure 7.7.2). M, marker-1kb ladder (Gibco BRL); Ov, ovaries; L, larvae (whole organism) used as a control. The primer sets generated the following products from transcribed larvae RNA: 974 bp, 780 bp (Z1); 320bp (Z2); 784bp (Z3); 1082 bp (Z4). The PCR did not generate the 728bp Z1 product from transcribed ovarian RNA indicating the lack of this transcript during oogenesis.

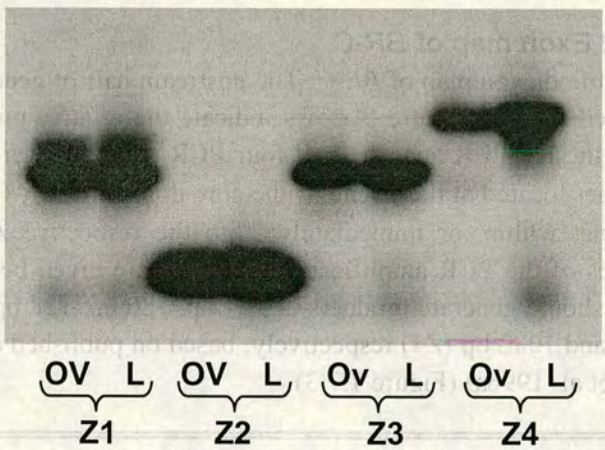


Figure 7.7.4 Southern analysis of *BR-C* transcripts

The gel from the RT-PCR analysis of the *BR-C* transcripts was transferred to a nylon membrane and hybridised with a radio labelled probe for the *BR-C* core domain. The probe was prepared from the Bs-cD5 plasmid DNA provided by (Bayer et al. 1996a) (for details about the plasmid see Table 2.2.1; for probe preparation and hybridisation techniques, 2.5.13.2 chapter.

It seems likely, therefore, that, as in metamorphosis, all zinc finger isoforms are expressed and function to regulate downstream gene expression. However, only Z1 and partially Z4 are expressed at a high enough levels to be detected by *in situ* and immunohystochemistry methods.

7.8 *Ectopic* BR-C *expression during mid-oogenesis affects endoreplication and chorion gene amplification*

Ectopic *BR-C* expression appears to induce premature production of the chorion. Figures 7.8.1C and 7.8.1D show that the chorion is already present in the stage 11 egg chamber. This could isolate the oocyte from the nurse cells and physically prevent transferring of the nurse cell components into the oocyte. Alternatively this could result from an altered pattern of transcription and translation of the chorion genes, or from abnormalities in the amplification of the chorion genes, or both. We investigated, therefore, whether the alterations in *BR-C* expression affected the timing or pattern of chorion gene amplification. Since the chorion is synthesised by most follicle cells this function could be related to the earlier expression of the *BR*-C. To monitor amplification we investigated the incorporation of BrdU in the follicle cell nuclei of wild type ovaries and in ovaries misexpressing various isoforms of the *BR-C* (see methods,2.5.16) .

In wild type ovaries, after 8 mitotic cell divisions, the endoreplication phase of follicle cell development begins (stage 6 of oogenesis) and is completed by stage 10B; during this process the entire nucleus is labelled by BrdU. The endoreplication is asynchronous in wild type and w^{1118} (the host strain used to produce the transgenic lines) ovaries and occurs in both nurse cells and follicle cells (Figure 7.8.2A, C). We observed a continuous endoreplication in the nurse cell associated follicle cells at stage 10B (Figure 7.8.2C). This is followed by the chorion gene amplification phase when 4 spots of incorporation are seen per nucleus in the follicle cells overlying the oocyte (Figure 7.8.2C, G, K). These 4 spots represent amplification of the two clusters of chorion genes (Figure 7.8.2G, K). Two are always larger, presumably due to the higher level of amplification of the cluster on chromosome 3 compared to the X chromosome cluster (Orr-Weaver and Spradling 1986; Delidakis and Kafatos

1989). This amplification was first observed at the border between the oocyte and nurse cells and it soon spread to the rest of the follicle cells.

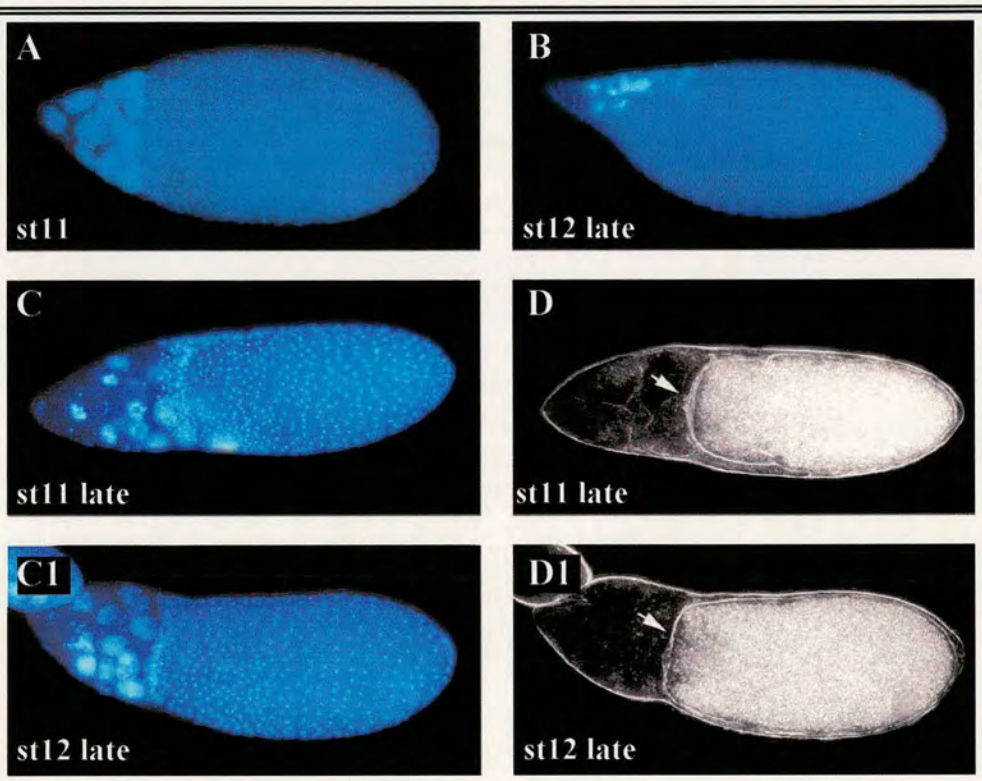


Figure 7.8.1 Ectopic *BR-C* expression affects the onset of chorion synthesis

A, D: Hoechst staining; C1–D1, dark field microscopy

A: Shows a wild type egg chamber at late stage 11. The oocyte is larger than the nurse cell complex due to the onset of nurse cell dumping into the oocyte.

B: Wild type egg chamber, late stage 12. Dumping is complete.

C-D: (late stage 11) and

C1-D1: (late stage 12) show that transgenic Z1-Z4 also causes inappropriate chorion synthesis (arrows), and this blocks dumping of the nurse cell cytoplasm into the oocyte during stage 11-12 of oogenesis.

When the *BR-C* isoforms are misexpressed, there is prolonged and synchronised endoreplication until late stage 10B, followed by specific amplification of genes in each nucleus Figure 7.8.2D, F, H, L). These results are observed from 3.5 – 4.5 hours after heatshock. We also observe extra spots of incorporated BrdU in the nuclei (Figure 7.8.2H, L).

There were three possible explanations for this: either the heatshock could be responsible; or the homologues of the chromosomes could have separated due to a defect in the cell cycle; or there could be amplification of DNA at additional sites in the genome. The host flies used to produce the transgenic lines, w^{1118}, were heatshocked and still showed 4 spots per nucleus, so heatshock itself was not responsible for the results. We counted the number of spots per nucleus and found 6 or 12 spots in approximately 80% of the nuclei. Occasionally we observed up to 28 spots. If the cell cycle was affected, and the homologues had separated, we would expect to see many more than 28 spots per nucleus due to the polyploidy of the follicle cells. If the amplification sites varied we could not predict the numbers, and indeed it may well be variable. This suggests that there are other sites in the genome induced to replicate by *BR-C* overexpression. We conclude that the endoreplication of DNA and the amplification of the chorion genes depends upon the *BR-C* encoded proteins or an unknown protein that is encoded by one of the downstream targets of the *BR-C*. This observation is consistent with the report that a mutation in the *BR-C* locus causes premature arrest of chorion gene amplification (Huang and Orr 1992).

Figure 7.8.2 Chorion gene amplification. BrdU incorporation associated with endoreplication and amplification

A: BrdU incorporation in the nuclei of wild type Drosophila *(Oregon R)* ovaries and w^{1118}, the host strain used for the production of the *BR-C* transgenic lines, stage 10A. The endocycles are not synchronous. Thus, just some of the nuclei are positive (Calvi et al. 1998).

B: Overexpression of a *BR-C* (Z1) transgene in ovaries at stage 10A. Most of the nuclei show synchronous endoreplication.

C: Normal synchronised amplification in wild type and w^{1118} at stage 10B. By that time the endoreplication associated with the main body follicle cells is completed.

D: BrdU incorporation in the nuclei of heatshocked *BR-C* (Z1) transgenic ovaries, stage 10B. Strong BrdU incorporation is present in the nurse cells.

E: BrdU incorporation in wild type nuclei at stage 12.

F: BrdU incorporation in the nuclei of *BR-C* (Z1) expressing ovaries 3h after the heatshock, stage 12. Some of the nuclei contain extra spots of replication.

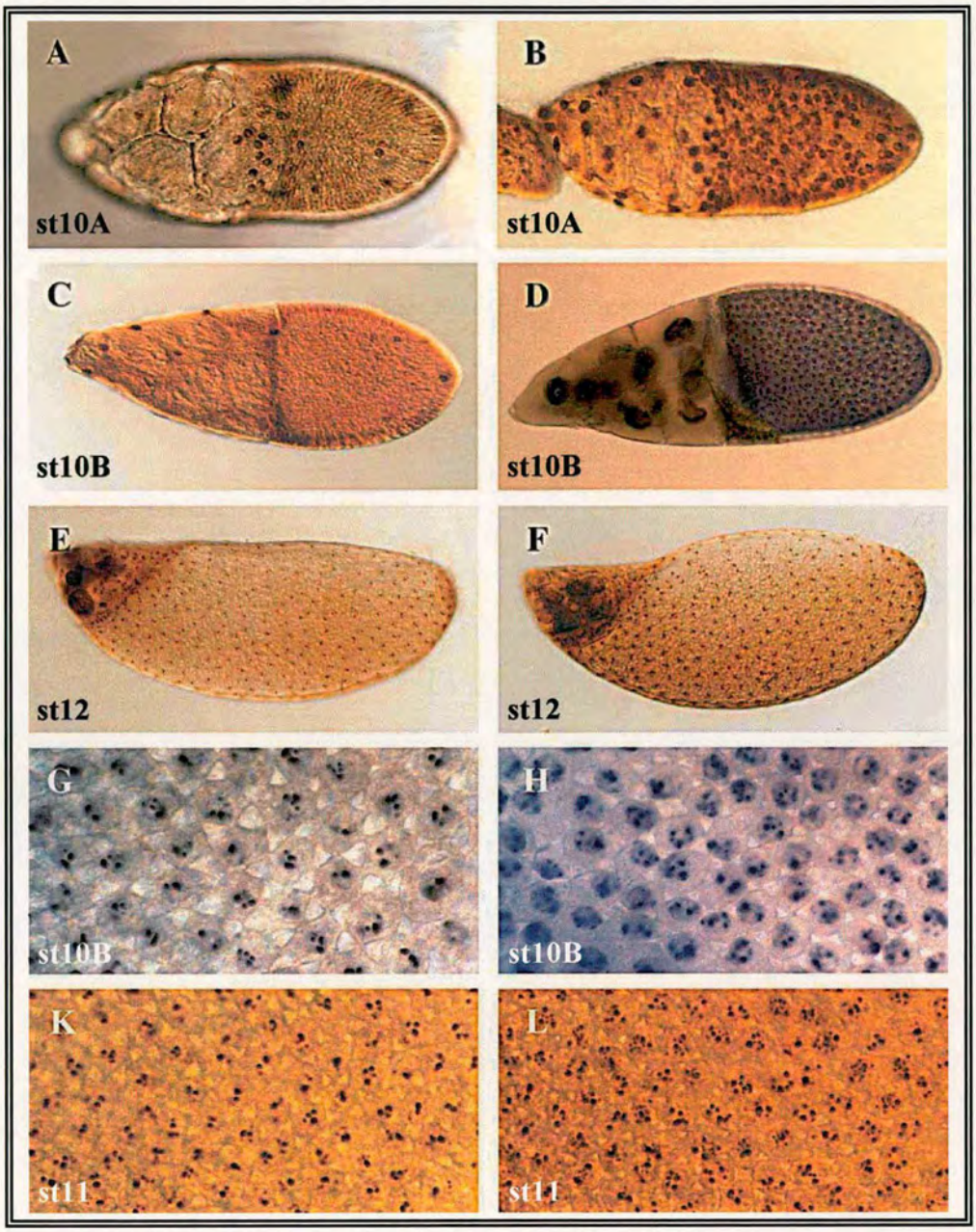
G: Higher magnification of wild type and w^{1118} nuclei at stage 10B to show the four spots of amplification normally occurring in wild type ovaries.

H: BrdU incorporation in the nuclei of *BR-C* (Z1) expressing ovaries 3h after the heatshock, stage 10B. Some of the nuclei contain extra spots of replication. The amplification pattern overlaps with the labelling due to continued endoreplication.

E-K: The nuclei of heatshocked w^{1118} ovaries still show four spots of amplification.

F-L: The nuclei of *BR-C* (Z1) transgenic ovaries 4.5 h after the heatshock exhibit a multispotted pattern.

Figure 7.8.2 Chorion gene amplification. BrdU incorporation associated with endoreplication and amplification



7.9 Ectopic BR-C expression in relation to chorion gene expression

The chorion is produced by the columnar follicle cells to provide a shell around the egg. Later in oogenesis, two groups of cells migrate anteriorly to produce the chorionic appendages and produce very large amounts of chorion material. Ectopic expression of the Z1 isoforms lead to chorionic appendage deposition by extra cells lying at the anterior of the egg filling in the mid-dorsal gap observed in wild type eggs. Often the follicle cells fail to migrate anteriorly over the remaining nurse cells at stage 11 and they are present, therefore, at a more posterior position. The pattern of *chorion* gene expression was compared in wild type ovaries and those expressing the Z1 isoform ectopically. In the wild type, we observe a high concentration of chorion transcripts in all follicle cells at stage 9, prior to their translation; they then become inactive and transcripts are again seen in stages 11-14 (Figure 7.9A, C, E, G). In the ovaries with ectopic *BR-C* gene expression, examined 3.5-4.5 hours after the heatshock, the same high concentration of chorion transcripts is observed in anterior follicle cells, even though their location in relation to the nurse cells is more posterior. This suggests the expression pattern is not dependent on the *BR-C* along the anterior-posterior axis (Figure 7.9B, D, F, H). Moreover, posterior follicle cells do not produce substantially more chorion material even though they express Z1 protein after heatshock. The fact that more dorsal cells produce appendage material and express the chorion genes means that *BR-C* expression in the most dorsal anterior cells does induce additional chorion production.

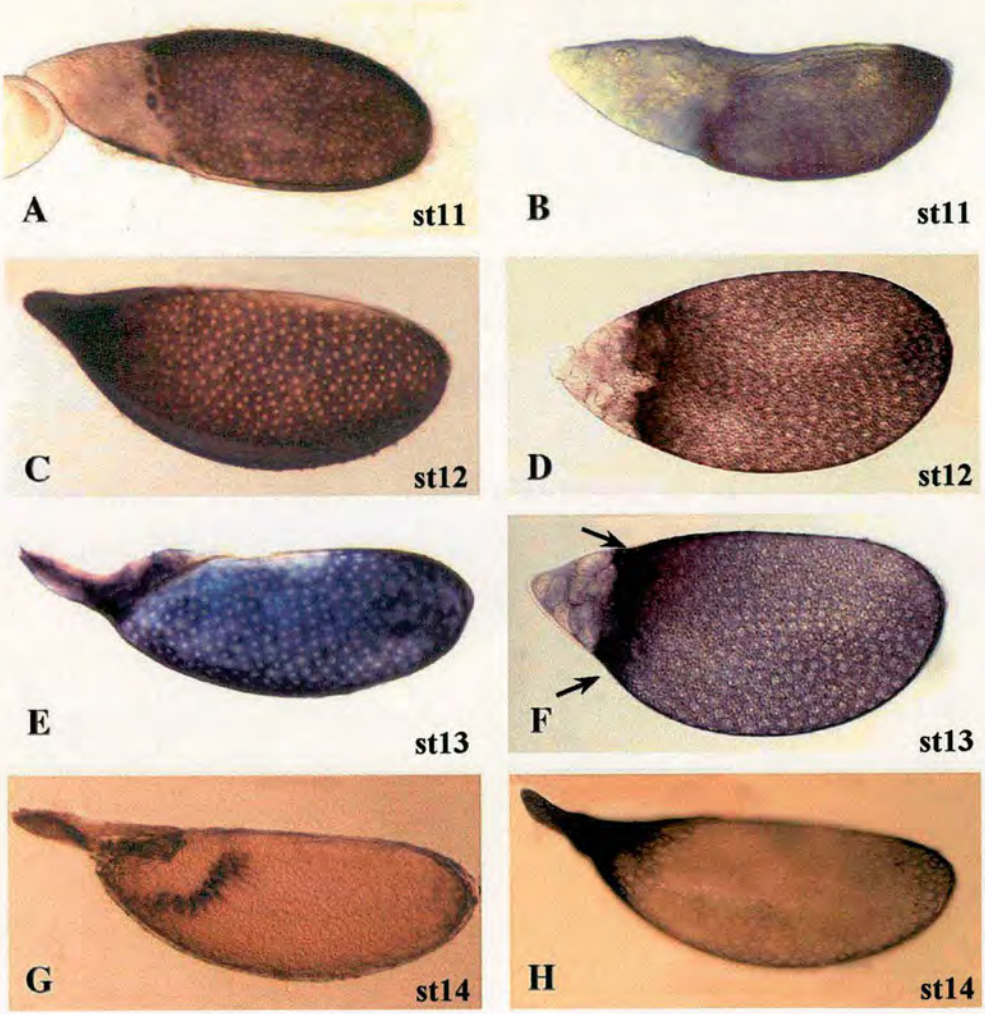


Figure 7.9 *In situ* hybridisation to chorion RNA in ovaries

A, C, E, G: *In situ* hybridisation to wild type (Oregon R) Drosophila ovaries with a DIG-labelled probe hybridising to chorion mRNA. The chorion probe, 7C8, was generated from cDNA encoding the c38 chorion protein, which maps to the X chromosome (see plasmid stocks, Table 2.2.1). Similar results were obtained using the alternative 7B7 probe (cDNA encoding c15 protein), which maps to a site on the third chromosome. For more information about the probes see Spradling *et al.* 1980.

B: *In situ* hybridisation to heatshocked Drosophila *BR-C* (Z1) transgenic ovaries with 7C8 probe, stage 11. At that stage an increase in the amount of chorion mRNA is observed at the border between the nurse cells and the oocyte. The misexpression of the *BR-C* leads to abnormal migration of the anterior follicle cells.

D: *In situ* pattern of chorion mRNA expression in heatshocked *BR-C* (Z1) transgenic ovaries, stage 12. A gap (arrows) is formed at the anterior pole as a result of failed migration of the follicle cells.

F: Strong expression of the chorion gene clusters in heatshocked *BR-C* (Z1) transgenic ovaries at stage 13. The variations in the egg chamber shape result from the heatshock induced misexpression of *BR-C*.

H: Strong continuous expression of chorion mRNA in heatshocked *BR-C* (Z1) transgenic ovaries, stage14. The anterior gap is not present because the nurse cells have completed dumping and the follicle cells have migrated correctly by the time of heatshock.

In summary we observed two different effects. Initially the ectopic expression of *BR-C* prevents the migration of the dorsal follicle cells in an anterior direction. Then the midline cells which normally express chorion protein at that stage start depositing chorion material in the wrong location. This results in the production of aberrant dorsal appendages. It is possible that the *BR-C* activates downstream genes which, in turn, activate the chorion genes. In anterior cells high levels of *BR-C* expression "win" over trans-acting repressors, but in posterior cells they do not. The *in situ* hybridisation results and the observed characteristic phenotypes following ectopic *BR-C* expression are both consistent with this observation. Alternatively *BR-C* could be essential for the cell migrations to position the follicle cells and endoreplication of the chorion genes but not in regulating chorion gene expression.

7.10 DISCUSSION

7.10.1 The *BR-C* complementing groups and zinc-finger isoforms

It has been shown by genetic analysis that rbp^+ function is required for dorsal appendage formation, and it was observed that Z1 is the only zinc-finger isoform expressed at high levels in the appendage-producing cells. These observations are compatible with the report that Z1 provides rbp^+ function (Bayer et al. 1997). It was also found that rbp^+ function is partially provided by Z4, but is not provided by Z2 and Z3, (Bayer et al. 1997). However, heatshock-induced expression of all four zinc-finger isoforms (Z1, Z2, Z3 and Z4) leads to a similar phenotype of extra dorsal appendage material production in the dorsal-gap, indicating that they may substitute for each other functionally in dorsal appendage formation.

The homozygous viable mutant, br^1, was the first mutant identified in the *BR-C* locus. br^1 mutant flies have shorter and wider wings than in wild type and fail to complement other mutations which belong to the *br* complementation group. However, the results from the complementation analysis suggests that the br^1 mutations also partially remove rbp^+ function. Therefore, it is in fact a *2Bab* allele (Figure 7.1.1). It is known that *2Bab* mutations cause reduction of both Z1 and Z2

expression. Thus in the br^1/br^{A47} and br^1/npr^6 females, both Z1 and Z2 are reduced. The reduction of Z1 levels results in the reduction of dorsal appendages, while no effect is produced by the reduction of Z2 levels. That is why no defects were observed in eggs produced by br^1/br^5 and br^1/br^6 females.

The mutant phenotypes clearly show the need for the *BR-C* in dorsal appendage morphogenesis. PCR experiments have demonstrated that all zinc finger isoforms are expressed in oogenesis, but as yet we have no evidence that they perform different functions. We also do not know the spatial and temporal distribution of Z2-Z4, which are not present at sufficiently high levels for detection by *in situ* hybridisation. Overexpression studies using transgenic flies carrying heatshock-controlled Z1, Z2, Z3, and Z4 isoforms lead to a failure of proper migration of the follicle cells that will secrete the appendages, premature chorion deposition, and abnormal appendage formation.

7.10.2 Relationship between the *BR-C* and chorion production

We have shown here that the *BR-C* is important for controlled DNA replication in oogenesis. Overexpression does not affect the timing of the onset of endoreplication and amplification but, endoreplication is prolonged beyond that observed in wild type ovaries and it leads to additional replication loci in the genome. These additional sites presumably share sequence similarities with the cis-acting sites regulating chorion gene amplification. This suggests that the *BR-C* is a key regulator of endoreplication and chorion gene amplification. The early *BR-C* expression pattern is in all the follicle cells and it is presumably at this stage that it is involved in this function. The expression of the *BR-C* is first observed in wild type flies at stage 6 and it is also at stage 6 that the endoreplication cycles begin. Since we do not observe premature endoreplication with *BR-C* overexpression, presumably other components essential for endoreplication are absent until stage 6 of oogenesis. The active role of the *BR-C* in endoreplication is also apparent from the fact that we observed prolonged incorporation of BrdU in the nurse cell nuclei when the *BR-C* is overexpressed. This presumably results in expression of the *BR-C* in the nurse cells, where it normally is not expressed. This shows that the proteins encoded by the *BR-C*

can function to prolong replication of DNA, even in cells where it is not normally used to control this process. It also suggests that an alternative regulator for DNA replication to the *BR-C* is used in the nurse cells. In normal development, the later *BR-C* expression, which is maintained in the dorsal-anterior cells making the appendages, is probably needed for cell migration, chorion deposition, and other follicle cell differentiation events.

It is possible that the initial activation of the *BR-C* in all follicle cells is regulated by the ecdysone/USP heterodimer (Yao et al. 1992; Horner et al. 1995) as is observed in metamorphosis. There is some evidence that in intermoult puffs a second heterodimer with DHR38 (a nuclear receptor related to NGF1-B from mammals) can compete for the same binding sites (Crispi et al. 1998). It is also possible that other DNA binding proteins besides the ecdysone receptor will be involved in *BR-C* expression. This would suggest a role for ecdysone during these stages of oogenesis. Later, the *BR-C* could be repressed in specific follicle cells by competitive and co-operative interactions with other gene products initiated by the *grk* and *dpp* signalling pathways. There is a significant amount of evidence that ecdysone and Juvenile Hormone (which binds to USP) are important for the progress of oogenesis (Bownes 1989) and we have recently shown that there is a check point in oogenesis that regulates whether egg chambers will proceed with development or undergo an apoptosis, which is regulated by the balance of Juvenile Hormone and ecdysone (Soller et al. 1999). However, there is little evidence as to precisely what is the role these hormones in regulating oocyte development and egg chamber differentiation. It has been shown (by antibody staining) that the Ecdysone Receptor is present in the follicle cells at the time when *BR-C* is activated (Mauchline, Deng and Bownes, unpublished).

Once activated, as we have shown, the *BR-C* gene is involved in endoreplication, the selective amplification of the chorion genes, and in the subsequent morphogenesis of the chorionic appendages. Calvi et al. (1998) have recently shown that the selective amplification of the chorion genes is closely linked with the cell cycle and the cycles of endoreplication that occur earlier in the follicle

cells. Somehow the chorion genes escape the replication controls that influence other parts of the genome. Our BrdU labelling experiments confirm their results on the timing of endoreplication and chorion amplification and the close association between endoreplication and selective amplification. Using overexpression of the *BR-C* we see not only the two extra sites they mentioned that may represent another chorion gene amplified for a function in later oogenesis, but also a number of additional sites. These may be sites with sequence similarity to the cis-acting sites regulating amplification. (Calvi et al. 1998) propose that there are amplification complexes located at chorion genes. Whether the *BR-C* encoded proteins are associated with these complexes or regulate the synthesis of one or more of their components remains to be elucidated. We have confirmed the link between endoreplication and chorion amplification and shown that it involves the *BR-C.* This may, therefore, provide a crucial link between hormones and the control of the cell cycle, and hence differentiation of the egg chamber during oogenesis.

In summary, our working model would be that the *BR-C* is activated by ecdysone in all follicle cells at stage 6 of oogenesis where its key function is the control of endoreplication, and then selective amplification. Later, when it is turned off in all but the anterior-dorsal follicle cells that will secrete the appendages, it has a second set of functions and is involved in the migration of cells and morphogenesis of the chorionic appendages. Recently this link between ecdysone (but not the ecdysone receptor), the *BR-C,* and morphogenesis has also been described for the progression of the furrow in the developing eye imaginal disc of *Drosophila* (Brennan et al. 1998; Brennan et al. 2001).

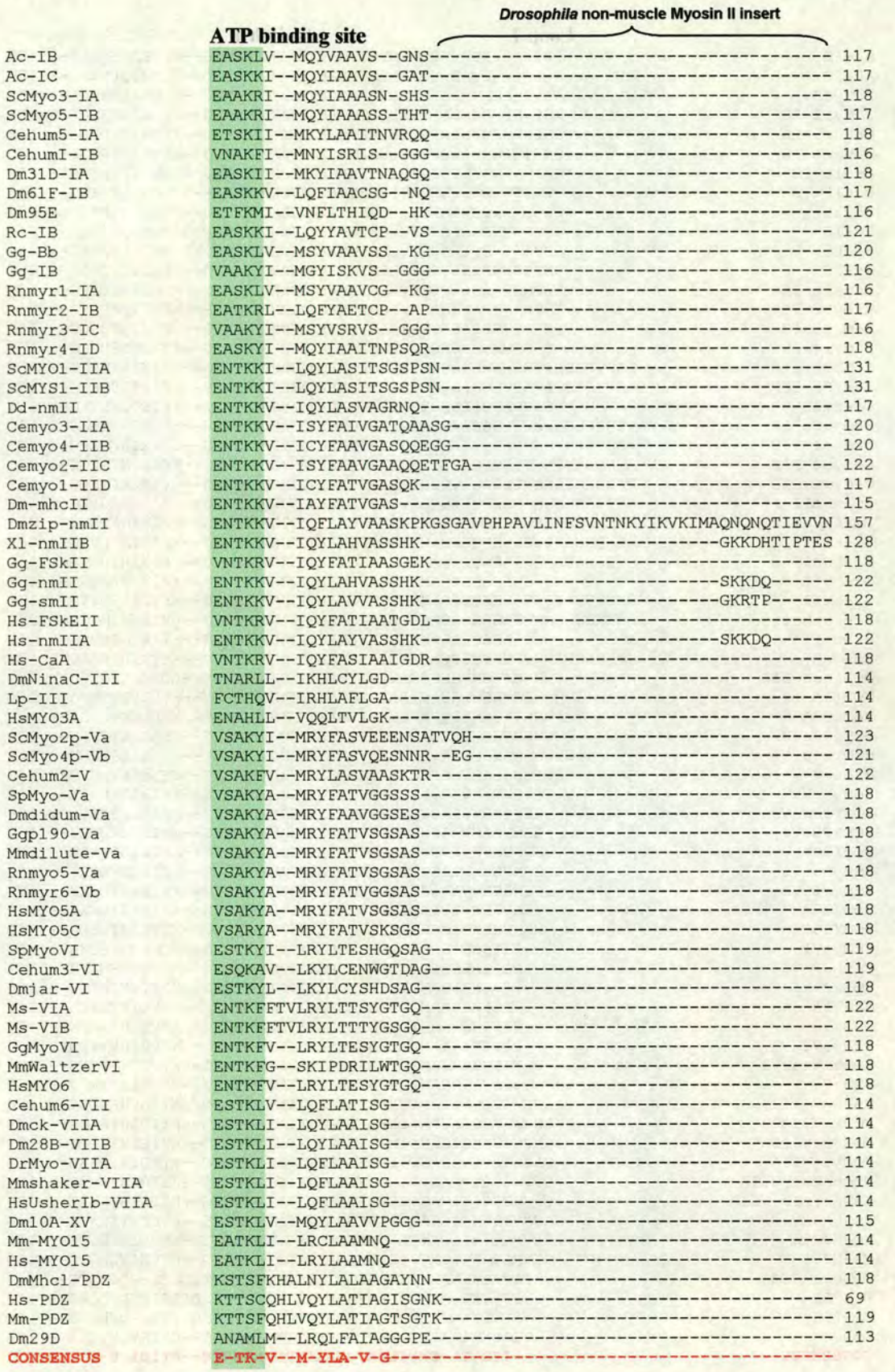
Appendix I Alignment of the conserved myosin head domains

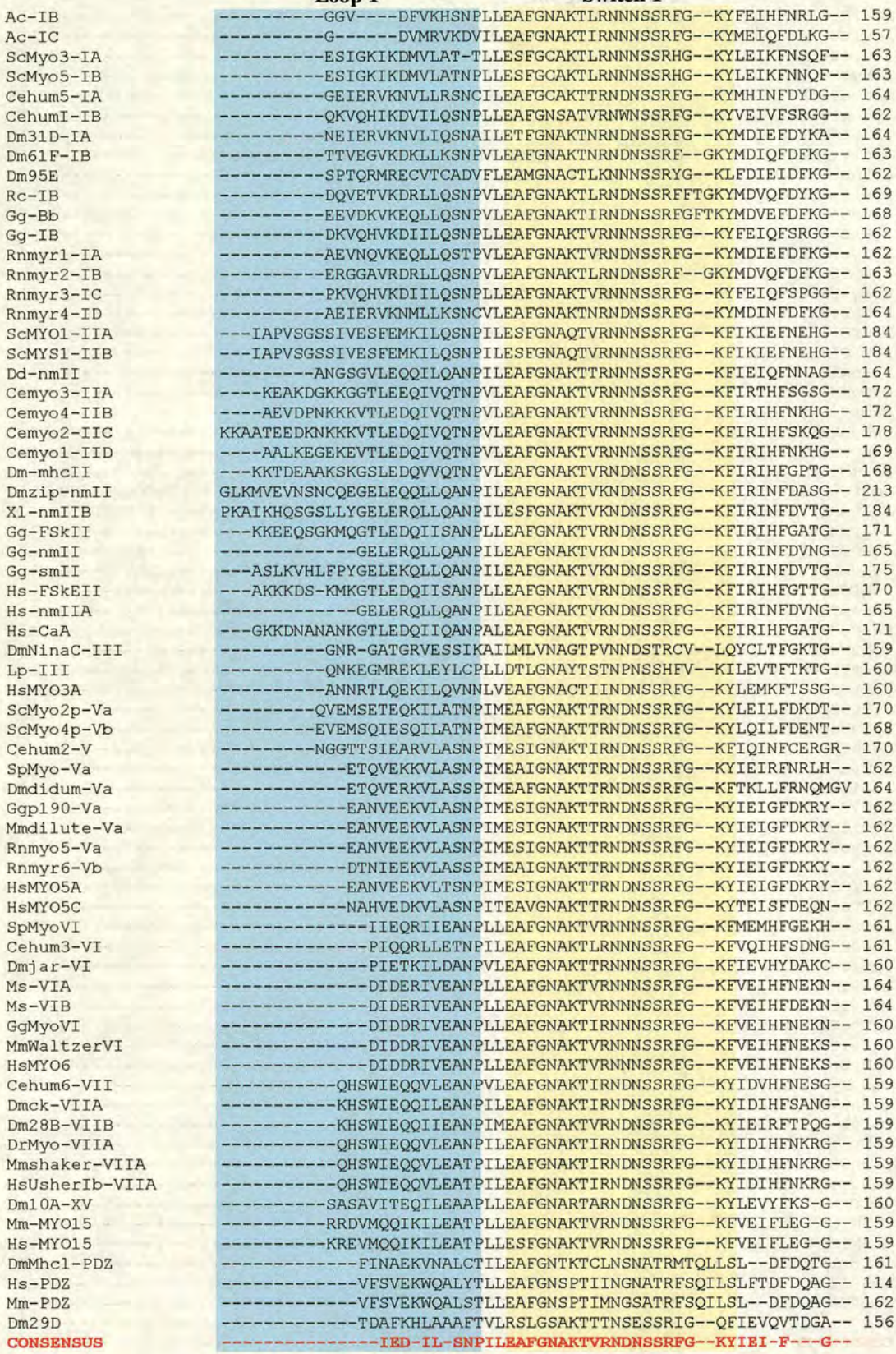
The head domains of *Drosophila* myosins (equivalent to amino acids 88-780 of chicken skeletal myosin) were compared to selected myosins representing phylogenetically diverse organisms. Some of the major structural features are colour shaded.

```
Ac-IB              VDDLVLMPKIT-EQDICANLEKRY-FNDLIYTNIG--PVLISVNPFRRIDA--LLTDECL 54
Ac-IC              VDDMVMLTSIS-NDAINDNLKKRF-AADLIYTYIG--HVLISVNPYKQINN--LYTERTL 54
ScMyo3-IA          ISDLTLLSKIS-DEAINENLKKRF-KNGIIYTYIG--HVLISVNPFRDLG---IYTNAVL 53
ScMyo5-IB          VSDLTLLSKIS-DEAINENLKKRF-LNATIYTYIG--HVLISVNPFRDLG---IYTDAVM 53
Cehum5-IA          VEDLVLLSTID-LKSVVQNLQLRF-QKGRIYTYIG--EVLVAVNPYRQLG---IYEKSTV 53
CehumI-IB          VDDMVLLPKLT-EQSIVENLKKRL-QANSIFTYIG--PVLISVNPFKQMP---YFTEKEM 53
Dm31D-IA           VQDFVLLDQVS-MEKFMDNLRKRF-QNGSIYTYIG--EVCVSMNPYRQMN---IYGPETI 53
Dm61F-IB           VQDFVLLENYQSEEAFIGNLKKRF-QEDLIYTYI--GQVLISVNPYKQLP---IYTDNHV 54
Dm95E              TWDSVLLENLS-EDSFINNIHQRY-KRDHIYTYIG--TSVVALNPYHHIS---EHSLDNV 53
Rc-IB              VQDFVLLENYTSEAAFIENLRKRF-KENLIYTYIFTGSVLVSVNPYKELE---IYSKQHM 56
Gg-Bb              VGDLVMLDPLS-EESLLRTLQERF-SRGEIYTYIGFTEVVISVNPYKPLP---IYTPEKV 55
Gg-IB              VDDMVLLAKIS-EEAIVENLKKRF-MDDYIFTYIG--PVLISVNPFKQMP---YFTDREI 53
Rnmyr1-IA          VGDTVLLEPLN-EETFIDNLKKRF-DHNEIYTYIG--SVVISVNPYRSLP---IYSPEKV 53
Rnmyr2-IB          VQDFVLLENFTSEAAFIENLRRRF-RENLIYTYI--GPVLVSVNPYRDLQ---IYTRQHM 54
Rnmyr3-IC          VDDMVLLSKIT-ESSIVENLKKRY-MDDYIFTYIG--SVLISVNPFKQMP---YFGEKEI 53
Rnmyr4-ID          KADFVLMDTVS-MPEFMANLRLRF-EKGRIYTFIG--EVVVSVNPYKVLN---IYGRDTI 53
ScMYO1-IIA         VENMSELTHFN-EPSVLYNLEKRY-DCDLIYTYSG--LFLVAINPYHNLN---LYSEDHI 53
ScMYS1-IIB         VENMSELTHLN-EPSVLYNLEKRY-DCDLIYTYSG--LFLVAINPYHNLN---LYSEDHI 53
Dd-nmII            VEDMSELSYLN-EPAVFHNLRVRY-NQDLIYTYSG--LFLVAVNPFKRIP---IYTQEMV 53
Cemyo3-IIA         TEDMANLTFLN-EASVLGNLKDRY-KDLMIYTYSG--LFCVVINPYKRLP---IYSESVI 53
Cemyo4-IIB         TEDMSNLSFLN-DASVLHNLRSRY-AAMLIYTYSG--LFCVVINPYKRLP---IYTDSCA 53
Cemyo2-IIC         TEDMSNLTFLN-DASVLYNLKARY-AAMLIYTYSG--LFCVVINPYKRLP---IYTESVA 53
Cemyo1-IID         TEDMSNLTFLN-DASVLWNLRSRY-AAMLIYTYSG--LFCVVINPYKRLP---IYTDSVA 53
Dm-mhcII           AEDMSNLTYLN-DASVLHNLRQRY-YNKLIYTYSG--LFCVAINPYKRYP---VYTNRCA 53
Dmzip-nmII         VEDMAELTCLN-EASVLHNIKDRY-YSGLIYTYSG--LFCVVVNPYKKLP---IYTEKIM 53
Xl-nmIIB           VEDMAELTCLN-EASVLHNLKDRY-YSGLIYTYSG--LFCVVINPYKNLP---IYSENII 53
Gg-FSkII           IEDMAMMTHLH-EPAVLYNLKERY-AAWMIYTYSG--LFCVTVNPYKWLP---VYNPEVV 53
Gg-nmII            VEDMAELTCLN-EASVLHNLKERY-YSGLIYTYSG--LFCVVINPYKNLP---IYSEEIV 53
Gg-smII            VEDMAELTCLN-EASVLHNLRERY-FSGLIYTYSG--LFCVVINPYKQLP---IYSEKII 53
Hs-FSkEII          IEDMAMLTHLN-EPAVLYNLKDRY-TSWMIYTYSG--LFCVTVNPYKWLP---VYNPEVV 53
Hs-nmIIA           VEDMAELTCLN-EASVLHNLKERY-YSGLIYTYSG--LFCVVINPYKNLP---IYSEEIV 53
Hs-CaA             IEDMAMLTFLH-EPAVLFNLKERY-AAWMIYTYSG--LFCVTVNPYKWLP---VYNAEVV 53
DmNinaC-III        PEDLAALENPV-DENIIESLRHRI-LMGESYSFIG--DILLSLNSNEIKQ---EFPQEFH 53
Lp-III             VDDLAFLDSPT-EEVVLENLEQRY-RKGEIYTFAG--DVLLTLNPGKVLP---LYGDQTA 53
HsMYO3A            VDDLATLEILD-ENTVSEQLEKCY-SRDQIYVYVG--DILIALNPFQSLG---LYSTKHS 53
ScMyo2p-Va         TEDLTSLSYLN-EPAVLHAIKQRYSQLN-IYTYSG--IVLIATNPFDRVDQ--LYTQDMI 54
ScMyo4p-Vb         TDDLTTLSYLN-EPAVLHAIKKRYMNGQ-IYTYSG--IVLIAANPFDKVDH--LYSREMI 54
Cehum2-V           KDDLTLLSYLH-EPAVLHNLQVRFVKGSSIYTYCG--IVLVAINPYADCSH--IYGEEII 55
SpMyo-Va           ENDLTSLSYLH-EPAVLYNLNERFIRNTAIYTYCG--IVLVAINPYESLP---LYGEDVI 54
Dmdidum-Va         QNDLTTLSYLH-EPGVLHNLRVRFCERQIIYTYCG--IILVAINPYAEMP---LYGPSII 54
Ggp190-Va          ENDLTALSYLH-EPAVLHNLKVRFIDSKLIYTYCG--IVLVAINPYEQLP---IYGEDII 54
Mmdilute-Va        ENDLTALSYLH-EPAVLHNLRVRFIDSKLIYTYCG--IVLVAINPYEQLP---IYGEDII 54
Rnmyo5-Va          ENDLTALSYLH-EPAVLHNLRVRFIDSKLIYTYCG--IVLVAINPYEQLP---IYGEDII 54
Rnmyr6-Vb          ENDLTALSHLH-EPAVLHNLKVRFLESNHIYTYCG--IVLVAINPYEQLP---IYGQDVI 54
HsMYO5A            ENDLTALSYLH-EPAVLHNLRVRFIDSKLIYTYCG--IVLVAINPYEQLP---IYGEDII 54
HsMYO5C            ENDLTALSYLH-EPAVLHNLRIRFAESKLIYTYSG--IILVAMNPYKQLP---IYGDAII 54
SpMyoVI            VDDNCSLMYLN-EATLLNNLRLRY-KKNSIYTYVA--NILIAVNPYKQMG--DLYTSSNI 54
Cehum3-VI          VEDNCALVHLN-EATLLNNCRLRY-ANGKIYTYVA--NILISINPYQLID--GLYSPETI 54
Dmjar-VI           HDDNCELMLLN-EATFLDNLKTRY-YKDKIYTYVA--NILIAVNPYREIK--ELYAPDTI 54
Ms-VIA             VEDNCSLMYLN-EATLLNNVRVRY-SKDKIYTFVA--NILIAVNPYYDIPFTKLYSPETI 56
Ms-VIB             VDDNCSLMYLN-EATLLNNVRVRY-SKDHIYTYVA--NILIAVNPYYDIPFTKLYGPDSI 56
GgMyoVI            VEDNCSLMYLN-EATLLHNIKVRY-SKDRIYTYVA--NILIAVNPYFDIP--KFYSSDAI 54
MmWaltzerVI        VEDNCSLMYLN-EATLLHNVKVRY-SKDRIYTYVA--NILIAVNPYFDIP--KIYSSDTI 54
HsMYO6             VEDNCSLMYLN-EATLLHNIKVRY-SKDRIYTYVA--NILIAVNPYFDIP--KIYSSEAI 54
Cehum6-VII         VEDMCQLGDFH-ESAILRNLFIRY-REKLIYAYTG--SILIAVNPYMDIA---IYTADEI 53
Dmck-VIIA          VEDMISLGDLH-EAGILRNLLIRY-KENLIYTYTG--SILVAVNPYQILP---IYTGDQI 53
Dm28B-VIIB         VEDMITLGDLQ-EYTILRNLQNRY-AKQLIYTYTG--SMLVAINPYQILP---IYTNREI 53
DrMyo-VIIA         VEDMIRLGDLN-EAGILRNLLIRY-REHLIYTYTG--SILVAVNPYQLLP---IYTADQI 53
Mmshaker-VIIA      VEDMIRLGDLN-EAGILRNRLIRY-RDHLIYTYTG--SILVAVNPYQLLS---IYSPEHI 53
HsUsherIb-VIIA     VEDMIRLGDLN-EAGILRNLLIRY-RDHLIYTYTG--SILVAVNPYQLLS---IYSPEHI 53
Dm10A-XV           VEDMTLLDDLH-EASLLWNLRLRY-DKGLIYTFAG--SILIAVNPYKMFPD--AYGLEVA 54
Mm-MYO15           VEDMTQLEDLQ-ETTVLANLKTRF-ERNLIYTYIG--SILVSVNPYRMFA---IYGPEQV 53
Hs-MYO15           VEDMTQLEDLQ-ETTVLSNLKIRF-ERNLIYTYIG--SILVSVNPYQMFG---IYGPEQV 53
DmMhc1-PDZ         VEDICELKYLN-EASVLHCLRQRY-ASNLIHTKAG--PTLLVVNPMAPLS---LYSEKVV 53
Hs-PDZ             ------------------------------------------------------------M 1
Mm-PDZ             LEDLASLVYLN-ESSVLHTLRQRY-GASLLHTYAG--PSLLVLSTRGAPA---VYSEKVM 53
Dm29D              -QDLIHLSGPLTEHAVMRTLQARF-NERRYFTNVG--PILLSINPYLDVG-----NPLTL 51
CONSENSUS          VEDM--L--LN-E-SVL-NLK-RY----LIYTY-G---VLV VNPYK-LP---IYS-D-I
```

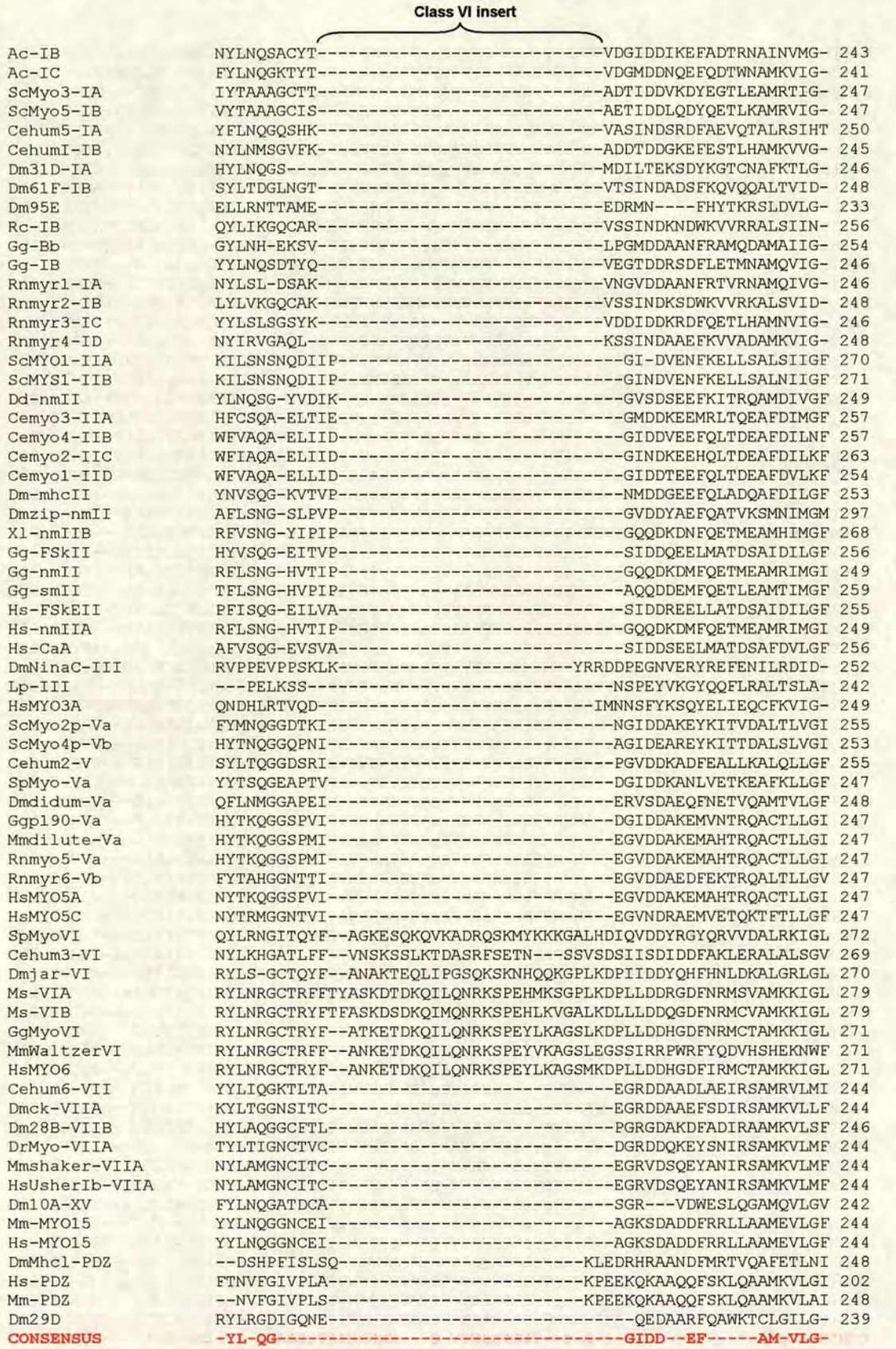
ATP binding site

```
Ac-IB              HCYR-GR----------YQHEQPPHVYALAEAAYRGVKSENINQCVIIS--GESGAGKT 100
Ac-IC              KDYR-GK----------YRYELPPHVYALADDMYRTMLSESEDQCVIIS--GESGAGKT 100
ScMyo3-IA          ESYK-RKK---------NRLEVPPHVFAIAESMYYNLKSYNENQCVIIS--GESGAGKT 100
ScMyo5-IB          NEYK-GK----------NRLEVPPHVFAIAESMYYNMKSYNENQCVIIS--GESGAGKT 99
Cehum5-IA          DQYK-GR----------EIYERAPHVFAIADAAYRSMKRFGRDSCIVIS--GESGAGKT 99
CehumI-IB          LLYQ-GA----------AQYGNAPHIYALADNMYRNMLIDNESQCVIIS--GESGAGKT 99
Dm31D-IA           RKYK-GR----------ELFENAPHLFAIADSAYRVLKQRQQDTCILIS--GESGAGKT 99
Dm61F-IB           KAYR-NK----------HFYEMPPHIFAVTDNAFRSLIEENRGQCVLIS--GESGSGKT 100
Dm95E              RNYG-DK----------GIFQLPPHIYGLTNLAYQSLKDQSEDQCVLLT--GESGAGKT 99
Rc-IB              ERYR-GV----------SFYEVSPHIYAIADNSYRSLRTERKDQCILISFTGESGAGKT 104
Gg-Bb              EEYH-NC----------NFFAVKPHIYAIADDAYRSLRDRDRDQCILITGFTESGAGKT 103
Gg-IB              ELYQ-GA----------AQYENPPHIYALTDNMYRNMLIDGENQCVIIS--GESGAGKT 99
Rnmyr1-IA          EDYR-NR----------NFYELSPHIFALSDEAYRSLRDQDKDQCILITG--ESGAGKT 99
Rnmyr2-IB          ERYR-GV----------SFYEVPPHLFAVADTVYRALRTERRDQAVMIS--GESGAGKT 100
Rnmyr3-IC          EMYQ-GA----------AQYENPPHIYALADSMYRNMIIDRENQCVIIS--GESGAGKT 99
Rnmyr4-ID          EQYK-GR----------ELYERPPHLFAILDAAYKAMKRRSKDTCIMIS--GESGAGKT 99
ScMYO1-IIA         NLYHNKHNRLSKSRLDENSHEKLPPHIFAIAEEAYENLLSEGKDQSILVT--GESGAGKT 111
ScMYS1-IIB         NLYHNKHNRLSKSRLDENSHEKLPPHIFAIAEEAYENLLSEGKDQSILVT--GESGAGKT 111
Dd-nmII            DIFKGRR----------RNEVAPHIFAISDVAYRSMLDDRQNQSLLIT--GESGAGKT 99
Cemyo3-IIA         KHFMGKR----------RNEMPPHLFAVSDEAYRNMVQDKENQSMLIT--GESGAGKT 99
Cemyo4-IIB         RMFMGKR----------KTEMPPHLFAVSDEAYRNMLQDHENQSMLIT--GESGAGKT 99
Cemyo2-IIC         RMFMGKR----------RTEMPPHLFAVSDEAYRNMLQNHENQSMLIT--GESGAGKT 99
Cemyo1-IID         RMFMGKR----------RTEMPPHLFAVSDQAYRYMLQDHENQSMLIT--GESGAGKT 99
Dm-mhcII           KMYRGKR----------RNEVPPHIFAISDGAYVDMLTNHVNQSMLIT--GESGAGKT 99
Dmzip-nmII         ERYKGIK----------RHEVPPHVFAITDSAYRNMLGDREDQSILCT--GESGAGKT 99
Xl-nmIIB           EMYRGKK----------RHEMPPHIYAISESAYRCMLQDREDQSILCT--GESGAGKT 99
Gg-FSkII           LAYRGKK----------RQEAPPHIFSISDNAYQFMLTDRENQSILIT--GESGAGKT 99
Gg-nmII            EMYKGKK----------RHEMPPHIYAITDTAYRSMMQDREDQSILCT--GESGAGKT 99
Gg-smII            DMYKGKK----------RHEMPPHIYAIADTAYRSMLQDREDQSILCT--GESGAGKT 99
Hs-FSkEII          EGYRGKK----------RQEAPPHIFSISDNAYQFMLTDRENQSILIT--GESGAGKT 99
Hs-nmIIA           EMYKGKK----------RHEMPPHIYAITDTAYRSMMQDREDQSILCT--GESGAGKT 99
Hs-CaA             AAYRGKK----------RSEAPPHIFSISDNAYQYMLTDRENQSILIT--GESGAGKT 99
DmNinaC-III        AKYR-FK----------SRSENQPHIFSVADIAYQDMLHHKEPQHIVLS--GESYSGKS 99
Lp-III             VKYC-ER----------GRSDNPPHVFAVADRAYQQMLHHKSPQAVILS--GVSGSGKS 99
HsMYO3A            KLYI-GS----------KRTASPPHIFAMADLGYQSIITYNSDQCIVIS--GESGAGKT 99
ScMyo2p-Va         QAYA--GK---------RRGELEPHLFAIAEEAYRLMKNDKQNQTIVVS--GESGAGKT 100
ScMyo4p-Vb         QNYS--SK---------RKDELEPHLFAIAEEAYRFMVHEKANQTVVVS--GESGAGKT 100
Cehum2-V           QVYRGAGK---------SAREMDPHIFAVAEEAHFDMGAFGKSQSIIVS--GESGAGKT 103
SpMyo-Va           QAYH--GQ---------DMGAMDPHIFAVAEEAFKRMARFEQDQSIIVS--GESGAGKT 100
Dmdidum-Va         RAYR--GH---------AMGDLEPHIFALAEEAYTKLERENCNLSIIVS--GESGAGKT 100
Ggp190-Va          NAYS--GQ---------NMGDMDPHIFAVAEEAYKQMARDERNQSIIVS--GESGAGKT 100
Mmdilute-Va        NAYS--GQ---------NMGDMDPHIFAVAEEAYKQMARDERNQSIIVS--GESGAGKT 100
Rnmyo5-Va          NAYS--GQ---------NMGDMDPHIFAVAEEAYKQMARDERNQSIIVS--GESGAGKT 100
Rnmyr6-Vb          YAYS--GQ---------NMGDMDPHIFAVAEEAYKQMARDEKNQSIIVS--GESGAGKT 100
HsMYO5A            NAYS--GQ---------NMGDMDPHIFAVAEEAYKQMARDERNQSIIVS--GESGAGKT 100
HsMYO5C            HAYS--GQ---------NMGDMDPHIFAVAEEAYKQMARNNRNQSIIVS--GESGAGKT 100
SpMyoVI            KKYQ--GK---------SLGTLPPHVFAIADKSYREMITSKQSQAIIVS--GESGAGKT 100
Cehum3-VI          KEYR--GK---------SLGQMEPHIFAIADKAYREMRRIKTSQSIIVS--GESGAGKT 100
Dmjar-VI           KKYN--GR---------SLGELPPHVFAIADKAIRDMRVYKLSQSIIVS--GESGAGKT 100
Ms-VIA             KQYR--GR---------SLGTLPPHVYAIADKAYRDMKVLKMSQSIIVS--GESGAGKT 102
Ms-VIB             KSYR--GK---------SLGTLPPHVFAIADKAYRDMKVLKMSQSIIVS--GESGAGKT 102
GgMyoVI            KKYQ--GR---------SLGTLPPHVFAIADKAYRDMKVLKMSQSIIVS--GESGAGKT 100
MmWaltzerVI        KSYQ--GK---------SLGTMPPHVFGIADKAFRDMKVLKMSQSIIVS--GESGAGKT 100
HsMYO6             KSYQ--GK---------SLGTRPPHVFAIADKAFRDMKVLKMSQSIIVS--GESGAGKT 100
Cehum6-VII         RMYK--RK---------RIGELPPHIFAIADNAYTNMRREKKNQSVIIS--GESGAGKT 99
Dmck-VIIA          KLYK--ER---------KIGELPPHIFAIGDNAYAHMKRYRQDQCIVIS--GESGAGKT 99
Dm28B-VIIB         QLYR--NK---------SLAELPPHIFAISDNAFQRLQRLKENQCVVIS--GESGAGKT 99
DrMyo-VIIA         RLYT--NK---------KIGEMPPHIFAIADNCYFNMQRNNKDQCCIIS--GESGAGKT 99
Mmshaker-VIIA      RQYT--NK---------KIGEMPPHIFAIADNCYFNMKRNNRDQCCIIS--GESGAGKT 99
HsUsherIb-VIIA     RQYT--NK---------KIGEMPPHIFAIADNCYFNMKRNSRDQCCIIS--GESGAGKT 99
Dm10A-XV           KQYA--GR---------PLGSLPPHLFAIGAAAHAALPSP---QVVVIS--GESGSGKT 97
Mm-MYO15           QQYS--GR---------ALGENPPHLFAIANLAFAKMLDAKQNQCVIIS--GESGSGKT 99
Hs-MYO15           QQYN--GR---------ALGENPPHLFAVANLAFAKMLDAKQNQCIIIS--GESGSGKT 99
DmMhcl-PDZ         SMFRGCKT---------EDMPPHVYSLAQTAYRSLVETRRDQSLIFM--GRSGSG-- 97
Hs-PDZ             HMFKGCRR---------EDMAPHIYAVAQTAYRAMLMSRQDQSIILL--GSSGSGFT 47
Mm-PDZ             HMFKGCRR---------EDMAPHIYAVAQTAYRAMLMSRQDQSIVLL--GSSGSG-- 97
Dm29D              TSTR-------------AMPLAPQLQKIVQEAVRQQSETGYPQAIILS--GTSGAGKT 94
CONSENSUS          --YR----------------EMPPHIFAIAD-AYR-M------QSIIIS--GESGAGKT
```


Drosophila non-muscle Myosin II insert
ATP binding site
Ac-IB EASKLV--MQYVAAVS--GNS---------- 117
Ac-IC EASKKI--MQYIAAVS--GAT---------- 117
ScMyo3-IA EAAKRI--MQYIAAASN-SHS---------- 118
ScMyo5-IB EAAKRI--MQYIAAASS-THT---------- 117
Cehum5-IA ETSKII--MKYLAAITNVRQQ---------- 118
CehumI-IB VNAKFI--MNYISRIS--GGG---------- 116
Dm31D-IA EASKII--MKYIAAVTNAQGQ---------- 118
Dm61F-IB EASKKV--LQFIAACSG--NQ---------- 117
Dm95E ETFKMI--VNFLTHIQD--HK---------- 116
Rc-IB EASKKI--LQYYAVTCP--VS---------- 121
Gg-Bb EASKLV--MSYVAAVSS--KG---------- 120
Gg-IB VAAKYI--MGYISKVS--GGG---------- 116
Rnmyr1-IA EASKLV--MSYVAAVCG--KG---------- 116
Rnmyr2-IB EATKRL--LQFYAETCP--AP---------- 117
Rnmyr3-IC VAAKYI--MSYVSRVS--GGG---------- 116
Rnmyr4-ID EASKYI--MQYIAAITNPSQR---------- 118
ScMYO1-IIA ENTKKI--LQYLASITSGSPSN---------- 131
ScMYS1-IIB ENTKKI--LQYLASITSGSPSN---------- 131
Dd-nmII ENTKKV--IQYLASVAGRNQ---------- 117
Cemyo3-IIA ENTKKV--ISYFAIVGATQAASG---------- 120
Cemyo4-IIB ENTKKV--ICYFAAVGASQQEGG---------- 120
Cemyo2-IIC ENTKKV--ISYFAAVGAAQQETFGA---------- 122
Cemyo1-IID ENTKKV--ICYFATVGASQK---------- 117
Dm-mhcII ENTKKV--IAYFATVGAS---------- 115
Dmzip-nmII ENTKKV--IQFLAYVAASKPKGSGAVPHPAVLINFSVNTNKYIKVKIMAQNQNQTIEVVN 157
Xl-nmIIB ENTKKV--IQYLAHVASSHK----------GKKDHTIPTES 128
Gg-FSkII VNTKRV--IQYFATIAASGEK---------- 118
Gg-nmII ENTKKV--IQYLAHVASSHK----------SKKDQ------ 122
Gg-smII ENTKKV--IQYLAVVASSHK----------GKRTP------ 122
Hs-FSkEII VNTKRV--IQYFATIAATGDL---------- 118
Hs-nmIIA ENTKKV--IQYLAYVASSHK----------SKKDQ------ 122
Hs-CaA VNTKRV--IQYFASIAAIGDR---------- 118
DmNinaC-III TNARLL--IKHLCYLGD---------- 114
Lp-III FCTHQV--IRHLAFLGA---------- 114
HsMYO3A ENAHLL--VQQLTVLGK---------- 114
ScMyo2p-Va VSAKYI--MRYFASVEEENSATVQH---------- 123
ScMyo4p-Vb VSAKYI--MRYFASVQESNNR--EG---------- 121
Cehum2-V VSAKFV--MRYLASVAASKTR---------- 122
SpMyo-Va VSAKYA--MRYFATVGGSSS---------- 118
Dmdidum-Va VSAKYA--MRYFAAVGGSES---------- 118
Ggp190-Va VSAKYA--MRYFATVSGSAS---------- 118
Mmdilute-Va VSAKYA--MRYFATVSGSAS---------- 118
Rnmyo5-Va VSAKYA--MRYFATVSGSAS---------- 118
Rnmyr6-Vb VSAKYA--MRYFATVGGSAS---------- 118
HsMYO5A VSAKYA--MRYFATVSGSAS---------- 118
HsMYO5C VSARYA--MRYFATVSKSGS---------- 118
SpMyoVI ESTKYI--LRYLTESHGQSAG---------- 119
Cehum3-VI ESQKAV--LKYLCENWGTDAG---------- 119
Dmjar-VI ESTKYL--LKYLCYSHDSAG---------- 118
Ms-VIA ENTKFFTVLRYLTTSYGTGQ---------- 122
Ms-VIB ENTKFFTVLRYLTTTYGSGQ---------- 122
GgMyoVI ENTKFV--LRYLTESYGTGQ---------- 118
MmWaltzerVI ENTKFG--SKIPDRILWTGQ---------- 118
HsMYO6 ENTKFV--LRYLTESYGTGQ---------- 118
Cehum6-VII ESTKLV--LQFLATISG---------- 114
Dmck-VIIA ESTKLI--LQYLAAISG---------- 114
Dm28B-VIIB ESTKLI--LQYLAAISG---------- 114
DrMyo-VIIA ESTKLI--LQFLAAISG---------- 114
Mmshaker-VIIA ESTKLI--LQFLAAISG---------- 114
HsUsherIb-VIIA ESTKLI--LQFLAAISG---------- 114
Dm10A-XV ESTKLV--MQYLAAVVPGGG---------- 115
Mm-MYO15 EATKLI--LRCLAAMNQ---------- 114
Hs-MYO15 EATKLI--LRYLAAMNQ---------- 114
DmMhcl-PDZ KSTSFKHALNYLALAAGAYNN---------- 118
Hs-PDZ KTTSCQHLVQYLATIAGISGNK---------- 69
Mm-PDZ KTTSFQHLVQYLATIAGTSGTK---------- 119
Dm29D ANAMLM--LRQLFAIAGGGPE---------- 113
CONSENSUS E-TK-V--M-YLA-V-G----------

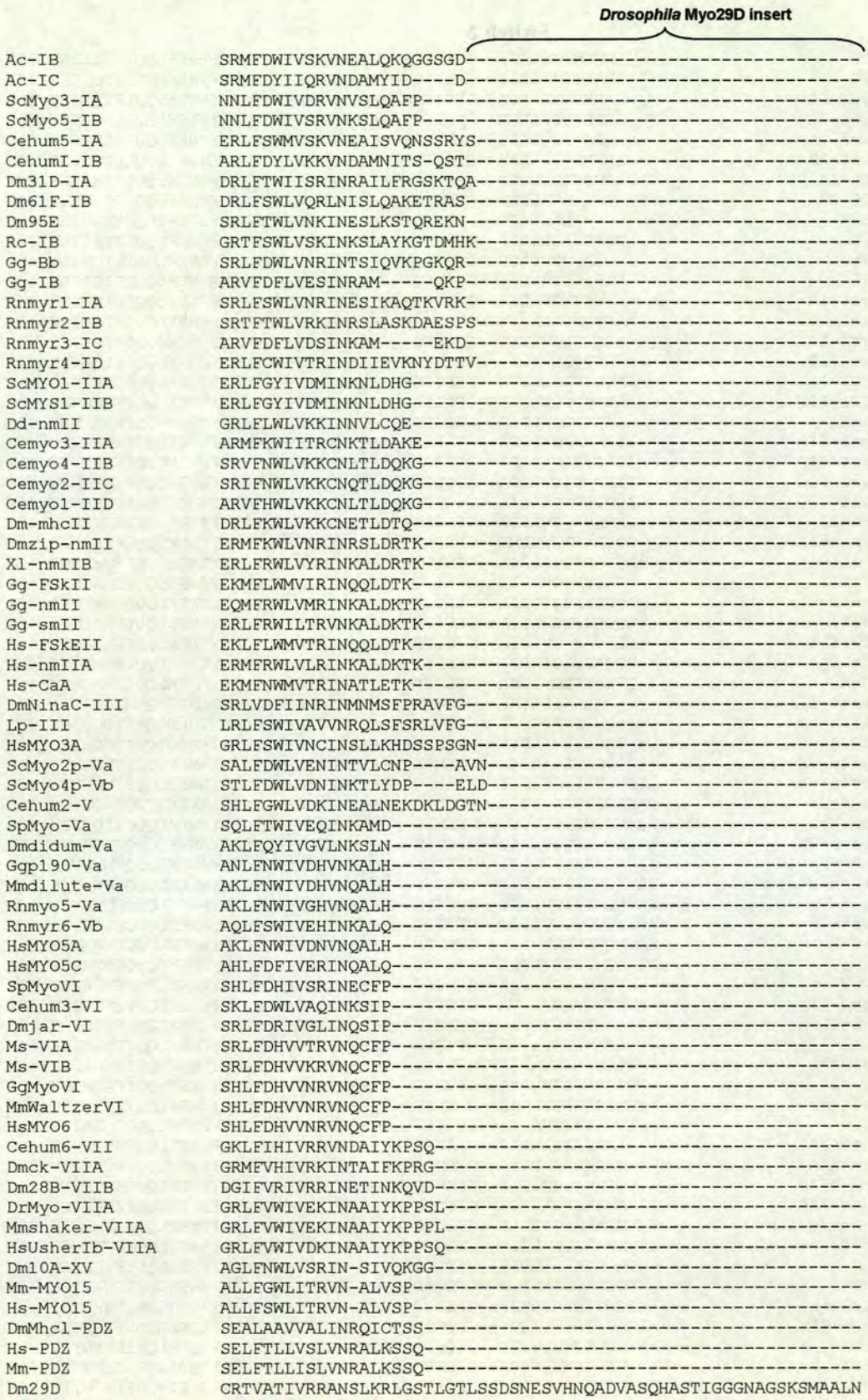

Loop 1
Switch 1
Ac-IB ----------GGV----DFVKHSNPLLEAFGNAKTLRNNNSSRFG--KYFEIHFNRLG-- 159
Ac-IC ----------G------DVMRVKDVILEAFGNAKTIRNNNSSRFG--KYMEIQFDLKG-- 157
ScMyo3-IA ----------ESIGKIKDMVLAT-TLLESFGCAKTLRNNNSSRHG--KYLEIKFNSQF-- 163
ScMyo5-IB ----------ESIGKIKDMVLATNPLLESFGCAKTLRNNNSSRHG--KYLEIKFNNQF-- 163
Cehum5-IA ----------GEIERVKNVLLRSNCILEAFGCAKTTRNDNSSRFG--KYMHINFDYDG-- 164
CehumI-IB ----------QKVQHIKDVILQSNPLLEAFGNSATVRNWNSSRFG--KYVEIVFSRGG-- 162
Dm31D-IA ----------NEIERVKNVLIQSNAILETFGNAKTNRNDNSSRFG--KYMDIEFDYKA-- 164
Dm61F-IB ----------TTVEGVKDKLLKSNPVLEAFGNAKTNRNDNSSRF--GKYMDIQFDFKG-- 163
Dm95E ----------SPTQRMRECVTCADVFLEAMGNACTLKNNNSSRYG--KLFDIEIDFKG-- 162
Rc-IB ----------DQVETVKDRLLQSNPVLEAFGNAKTLRNDNSSRFFTGKYMDVQFDYKG-- 169
Gg-Bb ----------EEVDKVKEQLLQSNPVLEAFGNAKTIRNDNSSRFGFTKYMDVEFDFKG-- 168
Gg-IB ----------DKVQHVKDIILQSNPLLEAFGNAKTVRNNNSSRFG--KYFEIQFSRGG-- 162
Rnmyr1-IA ----------AEVNQVKEQLLQSTPVLEAFGNAKTVRNDNSSRFG--KYMDIEFDFKG-- 162
Rnmyr2-IB ----------ERGGAVRDRLLQSNPVLEAFGNAKTLRNDNSSRF--GKYMDVQFDFKG-- 163
Rnmyr3-IC ----------PKVQHVKDIILQSNPLLEAFGNAKTVRNNNSSRFG--KYFEIQFSPGG-- 162
Rnmyr4-ID ----------AEIERVKNMLLKSNCVLEAFGNAKTNRNDNSSRFG--KYMDINFDFKG-- 164
ScMYO1-IIA ---IAPVSGSSIVESFEMKILQSNPILESFGNAQTVRNNNSSRFG--KFIKIEFNEHG-- 184
ScMYS1-IIB ---IAPVSGSSIVESFEMKILQSNPILESFGNAQTVRNNNSSRFG--KFIKIEFNEHG-- 184
Dd-nmII ---------ANGSGVLEQQILQANPILEAFGNAKTTRNNNSSRFG--KFIEIQFNNAG-- 164
Cemyo3-IIA ----KEAKDGKKGGTLEEQIVQTNPVLEAFGNAKTVRNNNSSRFG--KFIRTHFSGSG-- 172
Cemyo4-IIB ----AEVDPNKKKVTLEDQIVQTNPVLEAFGNAKTVRNNNSSRFG--KFIRIHFNKHG-- 172
Cemyo2-IIC KKAATEEDKNKKKVTLEDQIVQTNPVLEAFGNAKTVRNNNSSRFG--KFIRIHFSKQG-- 178
Cemyo1-IID ----AALKEGEKEVTLEDQIVQTNPVLEAFGNAKTVRNNNSSRFG--KFIRIHFNKHG-- 169
Dm-mhcII ---KKTDEAAKSKGSLEDQVVQTNPVLEAFGNAKTVRNDNSSRFG--KFIRIHFGPTG-- 168
Dmzip-nmII GLKMVEVNSNCQEGELEQQLLQANPILEAFGNAKTVKNDNSSRFG--KFIRINFDASG-- 213
Xl-nmIIB PKAIKHQSGSLLYGELERQLLQANPILESFGNAKTVKNDNSSRFG--KFIRINFDVTG-- 184
Gg-FSkII ---KKEEQSGKMQGTLEDQIISANPLLEAFGNAKTVRNDNSSRFG--KFIRIHFGATG-- 171
Gg-nmII ----------GELERQLLQANPILEAFGNAKTVKNDNSSRFG--KFIRINFDVNG-- 165
Gg-smII ---ASLKVHLFPYGELEKQLLQANPILEAFGNAKTVKNDNSSRFG--KFIRINFDVTG-- 175
Hs-FSkEII ---AKKKDS-KMKGTLEDQIISANPLLEAFGNAKTVRNDNSSRFG--KFIRIHFGTTG-- 170
Hs-nmIIA ----------GELERQLLQANPILEAFGNAKTVKNDNSSRFG--KFIRINFDVNG-- 165
Hs-CaA ---GKKDNANANKGTLEDQIIQANPALEAFGNAKTVRNDNSSRFG--KFIRIHFGATG-- 171
DmNinaC-III ----------GNR-GATGRVESSIKAILMLVNAGTPVNNDSTRCV--LQYCLTFGKTG-- 159
Lp-III ----------QNKEGMREKLEYLCPLLDTLGNAYTSTNPNSSHFV--KILEVTFTKTG-- 160
HsMYO3A ----------ANNRTLQEKILQVNNLVEAFGNACTIINDNSSRFG--KYLEMKFTSSG-- 160
ScMyo2p-Va ---------QVEMSETEQKILATNPIMEAFGNAKTTRNDNSSRFG--KYLEILFDKDT-- 170
ScMyo4p-Vb ---------EVEMSQIESQILATNPIMEAFGNAKTTRNDNSSRFG--KYLQILFDENT-- 168
Cehum2-V ---------NGGTTSIEARVLASNPIMESIGNAKTIRNDNSSRFG--KFIQINFCERGR- 170
SpMyo-Va ----------ETQVEKKVLASNPIMEAIGNAKTTRNDNSSRFG--KYIEIRFNRLH-- 162
Dmdidum-Va ----------ETQVERKVLASSPIMEAFGNAKTTRNDNSSRFG--KFTKLLFRNQMGV 164
Ggp190-Va ----------EANVEEKVLASNPIMESIGNAKTTRNDNSSRFG--KYIEIGFDKRY-- 162
Mmdilute-Va ----------EANVEEKVLASNPIMESIGNAKTTRNDNSSRFG--KYIEIGFDKRY-- 162
Rnmyo5-Va ----------EANVEEKVLASNPIMESIGNAKTTRNDNSSRFG--KYIEIGFDKRY-- 162
Rnmyr6-Vb ----------DTNIEEKVLASSPIMEAIGNAKTTRNDNSSRFG--KYIEIGFDKKY-- 162
HsMYO5A ----------EANVEEKVLTSNPIMESIGNAKTTRNDNSSRFG--KYIEIGFDKRY-- 162
HsMYO5C ----------NAHVEDKVLASNPITEAVGNAKTTRNDNSSRFG--KYTEISFDEQN-- 162
SpMyoVI ----------IIEQRIIEANPLLEAFGNAKTVRNNNSSRFG--KFMEMHFGEKH-- 161
Cehum3-VI ----------PIQQRLLETNPILEAFGNAKTLRNNNSSRFG--KFVQIHFSDNG-- 161
Dmjar-VI ----------PIETKILDANPVLEAFGNAKTTRNNNSSRFG--KFIEVHYDAKC-- 160
Ms-VIA ----------DIDERIVEANPLLEAFGNAKTVRNNNSSRFG--KFVEIHFNEKN-- 164
Ms-VIB ----------DIDERIVEANPLLEAFGNAKTVRNNNSSRFG--KFVEIHFDEKN-- 164
GgMyoVI ----------DIDDRIVEANPLLEAFGNAKTIRNNNSSRFG--KFVEIHFNEKN-- 160
MmWaltzerVI ----------DIDDRIVEANPLLEAFGNAKTVRNNNSSRFG--KFVEIHFNEKS-- 160
HsMYO6 ----------DIDDRIVEANPLLEAFGNAKTVRNNNSSRFG--KFVEIHFNEKS-- 160
Cehum6-VII ----------QHSWIEQQVLEANPVLEAFGNAKTIRNDNSSRFG--KYIDVHFNESG-- 159
Dmck-VIIA ----------KHSWIEQQILEANPILEAFGNAKTIRNDNSSRFG--KYIDIHFSANG-- 159
Dm28B-VIIB ----------KHSWIEQQIIEANPIMEAFGNAKTVRNDNSSRFG--KYIEIRFTPQG-- 159
DrMyo-VIIA ----------QHSWIEQQVLEANPILEAFGNAKTIRNDNSSRFG--KYIDIHFNKRG-- 159
Mmshaker-VIIA ----------QHSWIEQQVLEATPILEAFGNAKTIRNDNSSRFG--KYIDIHFNKRG-- 159
HsUsherIb-VIIA ----------QHSWIEQQVLEATPILEAFGNAKTIRNDNSSRFG--KYIDIHFNKRG-- 159
Dm10A-XV ----------SASAVITEQILEAAPLLEAFGNARTARNDNSSRFG--KYLEVYFKS-G-- 160
Mm-MYO15 ----------RRDVMQQIKILEATPLLEAFGNAKTVRNDNSSRFG--KFVEIFLEG-G-- 159
Hs-MYO15 ----------KREVMQQIKILEATPLLESFGNAKTVRNDNSSRFG--KFVEIFLEG-G-- 159
DmMhcl-PDZ ----------FINAEKVNALCTILEAFGNTKTCLNSNATRMTQLLSL--DFDQTG-- 161
Hs-PDZ ----------VFSVEKWQALYTLLEAFGNSPTIINGNATRFSQILSLFTDFDQAG-- 114
Mm-PDZ ----------VFSVEKWQALSTLLEAFGNSPTIMNGSATRFSQILSL--DFDQAG-- 162
Dm29D ----------TDAFKHLAAAFTVLRSLGSAKTTTNSESSRIG--QFIEVQVTDGA-- 156
CONSENSUS ----------IED-IL-SNPILEAFGNAKTVRNDNSSRFG--KYIEI-F---G--

```
Ac-IB                ---EPCGGRITNYLLEKSRVTFQTRGERSFHIFYQLLAGAS-DAEAQEMQLY---APENF 212
Ac-IC                ---DPVGGRISNYLLEKSRVVYQTNGERNFHIFYQLLAARA-RRPEAKFGLQ---TPDYY 210
ScMyo3-IA            ---EPCAGNITNYLLEKQRVVGQIKNERNFHIFYQFTKGAS-DTYKQMFGVQ---MPEQY 216
ScMyo5-IB            ---EPCAGNITNYLLEKQRVVSQIKNERNFHIFYQFTKGAS-DAYRQTFGVQ---KPEQY 216
Cehum5-IA            ---DPVGGNISNYLLEKSRVVRQQEGERNFHVFYQLVNGGD-DGLLRQFGLT--KDAKQY 218
CehumI-IB            ---EPIGGKLSNFLLEKSRVVHQNEGDRNFHVFYQLCAGD--KNLRSTFGIG---ELQYY 214
Dm31D-IA             ---DPVGGIITNYLLEKSRVVQQQPGERNFHSFYQLLRGAN-DNELRQYELQ--KETGKY 218
Dm61F-IB             ---APIGGNILNYLLEKSRVVAQMGGERNFHIFYQLLAGAD-EALLQELGLE--RALDTY 217
Dm95E                ---DPMG----------TRITDPIIGERNFHIFYQLLLGAD-LQLLKSLKLYR--NVEKY 206
Rc-IB                ---APVGGHILNYLLEKSRVVHQNHGERNFHIFYQLLEGGE-EDLLRRLGLFTDKNAQNY 225
Gg-Bb                ---DPLGGVISNYLLEKSRIVRHVKGERNFHIFYQLLAGGS-AQLLQQLKLRFTPDCSHY 224
Gg-IB                ---EPDGGKISNFLLEKSRVVSQNEHERNFHVYYQLIEGAS-QEQRQNLGIM---SPDYY 215
Rnmyr1-IA            ---DPLGGVISNYLLEKSRVVKQPRGERNFHVFYQLLSGAS-EELLHKLKLER--DFSRY 216
Rnmyr2-IB            ---APVGGHILSYLLEKSRVVHQNHGERNFHVFYQLLEGGE-EEALRRLGLE--RNPQSY 217
Rnmyr3-IC            ---EPDGGKISNFLLEKSRVVMRNPGERSFHIFYQLIEGAS-PEQKQSLGIT---SMDYY 215
Rnmyr4-ID            ---DPIGGHINNYLLEKSRVIVQQPGERSFHSFYQLLQGGS-EQMLHSLHLQ--KSLSSY 218
ScMYO1-IIA           ---MINGAHIEWYLLEKSRIVHQNSKERNYHIFYQLLSGLDDSELKNLRLKSRN--VKDY 239
ScMYS1-IIB           ---MINGAHIEWYLLEKSRIVHQNSKERNYHIFYQLLSGLDDSELKNLRLKSRN--VKDY 239
Dd-nmII              ---FISGASIQSYLLEKSRVVFQSETERNYHIFYQLLAGAT-AEEKKALHLAGP--ESFN 218
Cemyo3-IIA           ---KLAGGDIEHYLLEKSRVVRQAPGERCYHIFYQIMSGND-PSLRGKLKLSND--ITYY 226
Cemyo4-IIB           ---RLASCDIEHYLLEKSRVIRQAPGERCYHIFYQIYSDFR-PELKKELLLDLP--IKDY 226
Cemyo2-IIC           ---RVASCDIEHYLLEKSRVIRQAPGERSYHIFYQVFSDYL-PNLKKDLLLNKP--VKDY 232
Cemyo1-IID           ---TLASCDIEHYLLEKSRVIRQAPGERCYHIFYQIYSDFK-PQLRDELLLNHP--ISNY 223
Dm-mhcII             ---KLAGADIETYLLEKARVISQQSLERCYHIFYQIMSGSV-PGVKEMCFLSDN--IYDY 222
Dmzip-nmII           ---FISGANIETYLLEKSRAIRQAKDERTFHIFYQLLAGAT-PEQREKFILDD---VKSY 266
Xl-nmIIB             ---YIVGANIETYLLEKSRAVRQAKDERTFHIFYQLLAGSG-EHLKSDLLLDG---FNNY 237
Gg-FSkII             ---KLASADIETYLLEKSRVTFQLPAERSYHIFYQIMSNKK-PELIDMLLITTN--PYDY 225
Gg-nmII              ---YIVGANIETYLLEKSRAIRQAKEERTFHIFYYLLSGAG-EHLKTDLLLEP---YNKY 218
Gg-smII              ---YIVGANIETYLLEKSRAIRQAKDERTFHIFYYLIAGAS-EQMRNDLLLEG---FNNY 228
Hs-FSkEII            ---KLASADIETYLLEKSRVTFQLKAERSYHIFYQILSNKK-PELIELLLITTN--PYDY 224
Hs-nmIIA             ---YIVGANIETYLLEKSRAIRQAKEERTFHIFYYLLSGAG-EHLKTDLLLEP---YNKY 218
Hs-CaA               ---KLASADIETYLLEKSRVIFQLKAERNYHIFYQILSNKK-PELLDMLLVTNN--PYDY 225
DmNinaC-III          ---KMSGAVFNMYMLEKLRVATTDGTQHNFHIFYYFYDFINQQNQLKEYNLKADRNYRYL 216
Lp-III               ---KITGAILFTFLLEARRLTDIPKGERNFHVFYYFYEGLRSEGRLKEFGLEEK-NYRYL 216
HsMYO3A              ---AVVGAQISEYLLEKSRVIHQAIGEKNFHIFYYIYAGLAEKKKLAHYKLPENKPPRYL 217
ScMyo2p-Va           -SIIG--ARIRTYLLERSRLVYQPPIERNYHIFYQLMAGLP-AQTKEELHLT---DASDY 223
ScMyo4p-Vb           -TIRG--SKIRTYLLEKSRLVYQPETERNYHIFYQILEGLP-EPVKQELHLS---SPKDY 221
Cehum2-V             -RIVG--AEMKTYLLEKSRLVFQAPGERNYHIFYQLCAARN-HQVLKDLHLG---PCESY 223
SpMyo-Va             -HIVG--ANMRTYLLEKSRVVFQAPEERNYHIFYQLCACCD-QPELKELALG---HPDEF 215
Dmdidum-Va           MFLQG--ATMHTYLLEKSRVVYQAQGERNYHIFYQLCAAR---SKYPELVLD---HQDKF 216
Ggp190-Va            -RIIG--ANMRTYLLEKSRVVFQAEEERNYHIFYQLCASAA-LPEFKTLRLG---NANYF 215
Mmdilute-Va          -RIIG--ANMRTYLLEKSRVVFQAEEERNYHIFYQLCASAK-LPEFKMLRLG---NADSF 215
Rnmyo5-Va            -RIIG--ANMRTYLLEKSRVVFQAEEERNYHIFYQLCASAK-LPEFKMLRLG---NADSF 215
Rnmyr6-Vb            -HIIG--ANMRTYLLEKSRVVFQADDERNYHIFYQLCAAAS-LPEFKELALT---CAEDF 215
HsMYO5A              -RIIG--ANMRTYLLEKSRVVFQAEEERNYHIFYQLCASAK-LPEFKMLRLG---NADNF 215
HsMYO5C              -QIIG--ANMSTYLLEKSRVVFQSENERNYHIFYQLCASAQ-QSEFKHLKLG---SAEEF 215
SpMyoVI              -DVVG--GYVSHYLLEKPRVCWQNGEERNYHVFYYLCAGAP-DDLRTKLRIQ---TPDDY 214
Cehum3-VI            -TVAG--GFVSHYLLETSRVCRQAAGERNYHIFYQLIAGSS-PDLYKKLRLA---PASSF 214
Dmjar-VI             -QVVG--GYISHYLLEKSRICTQSAEERNYHVFYMLLAGAP-QQLRDKLSLG---KPDDY 213
Ms-VIA               -AVVGFTGFVSHYLLEKSRICRQSSEERNYHIFYRLCAGAT-EDIRQKFHLS---SPDTF 219
Ms-VIB               -AVVGFTGFVSHYLLEKSRICMQSNDERNYHIFYRLCAGAS-EDIKKKLHLD---SPDSF 219
GgMyoVI              -SVVG--GFVSHYLLEKSRICVQGKEERNYHIFYRLCAGAP-EDIREKLYLS---SPDSF 213
MmWaltzerVI          -SVVG--GFVSHYLLEKSRICVQGKEERNHHIFYRLCAGAS-EDIREKLHLS---SPDNF 213
HsMYO6               -SVVG--GFVSHYLLEKSRICVQGKEERNYHIFYRLCAGAS-EDIREKLHLS---SPDNF 213
Cehum6-VII           -SIEG--AKIEQYLLEKSRIVTQSENERNYHIFYCLLAGLS-REEKSELELG---TAADY 212
Dmck-VIIA            -VIEG--AKIEQYLLEKSRIVSQNHSERNYHVFYCILAGLS-ADEKSRLDLG---MAADY 212
Dm28B-VIIB           -AIQG--ARIQQYLLEKSRIVFQSRDERNYHIFYCMLAGLS-TAERERLKLQE-QSPSQY 214
DrMyo-VIIA           -AIEG--AKIEQYLLEKSRVCRQARDERNYHIFYCMLKGMT-PDQKKQLGLS---KATDY 212
Mmshaker-VIIA        -AIEG--AKIEQYLLSKSRVCRQAPDERNYHVFYCMLEGMN-EEEKKKLGLG---QAADY 212
HsUsherIb-VIIA       -AIEG--AKIEQYLLEKSRVCRQALDERNYHVFYCMLEGMS-EDQKKKLGLG---QASDY 212
Dm10A-XV             -AIVG--AKITQYLLEKSRIVTQAPGERNYHVFYELLGGLS-ETERSKYGLL---EADKY 213
Mm-MYO15             -VICG--AITSQYLLEKSRIVFQAKNERNYHIFYELLAGLP-AQLRQAFSLQ---EAETY 212
Hs-MYO15             -VISG--AITSQYLLEKSRIVFQAKNERNYHIFYELLAGLP-AQLRQAFSLQ---EAETY 212
DmMhcl-PDZ           ---QIASASLQVLLPERQRAGRRLGHEHSFHIMTRLLAGAA-GLLQKELHLE---NITSE 214
Hs-PDZ               ---QVASASIQTMLLEKLRVARRPASEATFNVFYYLLACGD-GTLRTELHLN---HLAEN 167
Mm-PDZ               ---QVASASIQTMLLEKLRVARRPASEATFNVFYYLLACGD-ATLRTELHLN---HLAEN 215
Dm29D                ----LYRTKIHCYFLDQTRVIRPLPKEKNYHIFYQLLAGLS-REERQKLHLDG-YSPANL 210
CONSENSUS            -------A-I--YLLEKSRVV-Q---ERNYHIFYQLLAGA----LK--L-L--------Y
```


Class VI insert
Ac-IB NYLNQSACYT----------VDGIDDIKEFADTRNAINVMG- 243
Ac-IC FYLNQGKTYT----------VDGMDDNQEFQDTWNAMKVIG- 241
ScMyo3-IA IYTAAAGCTT----------ADTIDDVKDYEGTLEAMRTIG- 247
ScMyo5-IB VYTAAAGCIS----------AETIDDLQDYQETLKAMRVIG- 247
Cehum5-IA YFLNQGQSHK----------VASINDSRDFAEVQTALRSIHT 250
CehumI-IB NYLNMSGVFK----------ADDTDDGKEFESTLHAMKVVG- 245
Dm31D-IA HYLNQGS----------MDILTEKSDYKGTCNAFKTLG- 246
Dm61F-IB SYLTDGLNGT----------VTSINDADSFKQVQQALTVID- 248
Dm95E ELLRNTTAME----------EDRMN----FHYTKRSLDVLG- 233
Rc-IB QYLIKGQCAR----------VSSINDKNDWKVVRRALSIIN- 256
Gg-Bb GYLNH-EKSV----------LPGMDDAANFRAMQDAMAIIG- 254
Gg-IB YYLNQSDTYQ----------VEGTDDRSDFLETMNAMQVIG- 246
Rnmyr1-IA NYLSL-DSAK----------VNGVDDAANFRTVRNAMQIVG- 246
Rnmyr2-IB LYLVKGQCAK----------VSSINDKSDWKVVRKALSVID- 248
Rnmyr3-IC YYLSLSGSYK----------VDDIDDKRDFQETLHAMNVIG- 246
Rnmyr4-ID NYIRVGAQL----------KSSINDAAEFKVVADAMKVIG- 248
ScMYO1-IIA KILSNSNQDIIP----------GI-DVENFKELLSALSIIGF 270
ScMYS1-IIB KILSNSNQDIIP----------GINDVENFKELLSALNIIGF 271
Dd-nmII YLNQSG-YVDIK----------GVSDSEEFKITRQAMDIVGF 249
Cemyo3-IIA HFCSQA-ELTIE----------GMDDKEEMRLTQEAFDIMGF 257
Cemyo4-IIB WFVAQA-ELIID----------GIDDVEEFQLTDEAFDILNF 257
Cemyo2-IIC WFIAQA-ELIID----------GINDKEEHQLTDEAFDILKF 263
Cemyo1-IID WFVAQA-ELLID----------GIDDTEEFQLTDEAFDVLKF 254
Dm-mhcII YNVSQG-KVTVP----------NMDDGEEFQLADQAFDILGF 253
Dmzip-nmII AFLSNG-SLPVP----------GVDDYAEFQATVKSMNIMGM 297
Xl-nmIIB RFVSNG-YIPIP----------GQQDKDNFQETMEAMHIMGF 268
Gg-FSkII HYVSQG-EITVP----------SIDDQEELMATDSAIDILGF 256
Gg-nmII RFLSNG-HVTIP----------GQQDKDMFQETMEAMRIMGI 249
Gg-smII TFLSNG-HVPIP----------AQQDDEMFQETLEAMTIMGF 259
Hs-FSkEII PFISQG-EILVA----------SIDDREELLATDSAIDILGF 255
Hs-nmIIA RFLSNG-HVTIP----------GQQDKDMFQETMEAMRIMGI 249
Hs-CaA AFVSQG-EVSVA----------SIDDSEELMATDSAFDVLGF 256
DmNinaC-III RVPPEVPPSKLK----------YRRDDPEGNVERYREFENILRDID- 252
Lp-III ---PELKSS----------NSPEYVKGYQQFLRALTSLA- 242
HsMYO3A QNDHLRTVQD----------IMNNSFYKSQYELIEQCFKVIG- 249
ScMyo2p-Va FYMNQGGDTKI----------NGIDDAKEYKITVDALTLVGI 255
ScMyo4p-Vb HYTNQGGQPNI----------AGIDEAREYKITTDALSLVGI 253
Cehum2-V SYLTQGGDSRI----------PGVDDKADFEALLKALQLLGF 255
SpMyo-Va YYTSQGEAPTV----------DGIDDKANLVETKEAFKLLGF 247
Dmdidum-Va QFLNMGGAPEI----------ERVSDAEQFNETVQAMTVLGF 248
Ggp190-Va HYTKQGGSPVI----------DGIDDAKEMVNTRQACTLLGI 247
Mmdilute-Va HYTKQGGSPMI----------EGVDDAKEMAHTRQACTLLGI 247
Rnmyo5-Va HYTKQGGSPMI----------EGVDDAKEMAHTRQACTLLGI 247
Rnmyr6-Vb FYTAHGGNTTI----------EGVDDAEDFEKTRQALTLLGV 247
HsMYO5A NYTKQGGSPVI----------EGVDDAKEMAHTRQACTLLGI 247
HsMYO5C NYTRMGGNTVI----------EGVNDRAEMVETQKTFTLLGF 247
SpMyoVI QYLRNGITQYF--AGKESQKQVKADRQSKMYKKKGALHDIQVDDYRGYQRVVDALRKIGL 272
Cehum3-VI NYLKHGATLFF--VNSKSSLKTDASRFSETN---SSVSDSIISDIDDFAKLERALALSGV 269
Dmjar-VI RYLS-GCTQYF--ANAKTEQLIPGSQKSKNHQQKGPLKDPIIDDYQHFHNLDKALGRLGL 270
Ms-VIA RYLNRGCTRFFTYASKDTDKQILQNRKSPEHMKSGPLKDPLLDDRGDFNRMSVAMKKIGL 279
Ms-VIB RYLNRGCTRYFTFASKDSDKQIMQNRKSPEHLKVGALKDLLLDDQGDFNRMCVAMKKIGL 279
GgMyoVI RYLNRGCTRYF--ATKETDKQILQNRKSPEYLKAGSLKDPLLDDHGDFNRMCTAMKKIGL 271
MmWaltzerVI RYLNRGCTRFF--ANKETDKQILQNRKSPEYVKAGSLEGSSIRRPWRFYQDVHSHEKNWF 271
HsMYO6 RYLNRGCTRYF--ANKETDKQILQNRKSPEYLKAGSMKDPLLDDHGDFIRMCTAMKKIGL 271
Cehum6-VII YYLIQGKTLTA----------EGRDDAADLAEIRSAMRVLMI 244
Dmck-VIIA KYLTGGNSITC----------EGRDDAAEFSDIRSAMKVLLF 244
Dm28B-VIIB HYLAQGGCFTL----------PGRGDAKDFADIRAAMKVLSF 246
DrMyo-VIIA TYLTIGNCTVC----------DGRDDQKEYSNIRSAMKVLMF 244
Mmshaker-VIIA NYLAMGNCITC----------EGRVDSQEYANIRSAMKVLMF 244
HsUsherIb-VIIA NYLAMGNCITC----------EGRVDSQEYANIRSAMKVLMF 244
Dm10A-XV FYLNQGATDCA----------SGR---VDWESLQGAMQVLGV 242
Mm-MYO15 YYLNQGGNCEI----------AGKSDADDFRRLLAAMEVLGF 244
Hs-MYO15 YYLNQGGNCEI----------AGKSDADDFRRLLAAMEVLGF 244
DmMhcl-PDZ --DSHPFISLSQ----------KLEDRHRAANDFMRTVQAFETLNI 248
Hs-PDZ FTNVFGIVPLA----------KPEEKQKAAQQFSKLQAAMKVLGI 202
Mm-PDZ --NVFGIVPLS----------KPEEKQKAAQQFSKLQAAMKVLAI 248
Dm29D RYLRGDIGQNE----------QEDAARFQAWKTCLGILG- 239
CONSENSUS -YL-QG----------GIDD--EF-----AM-VLG-

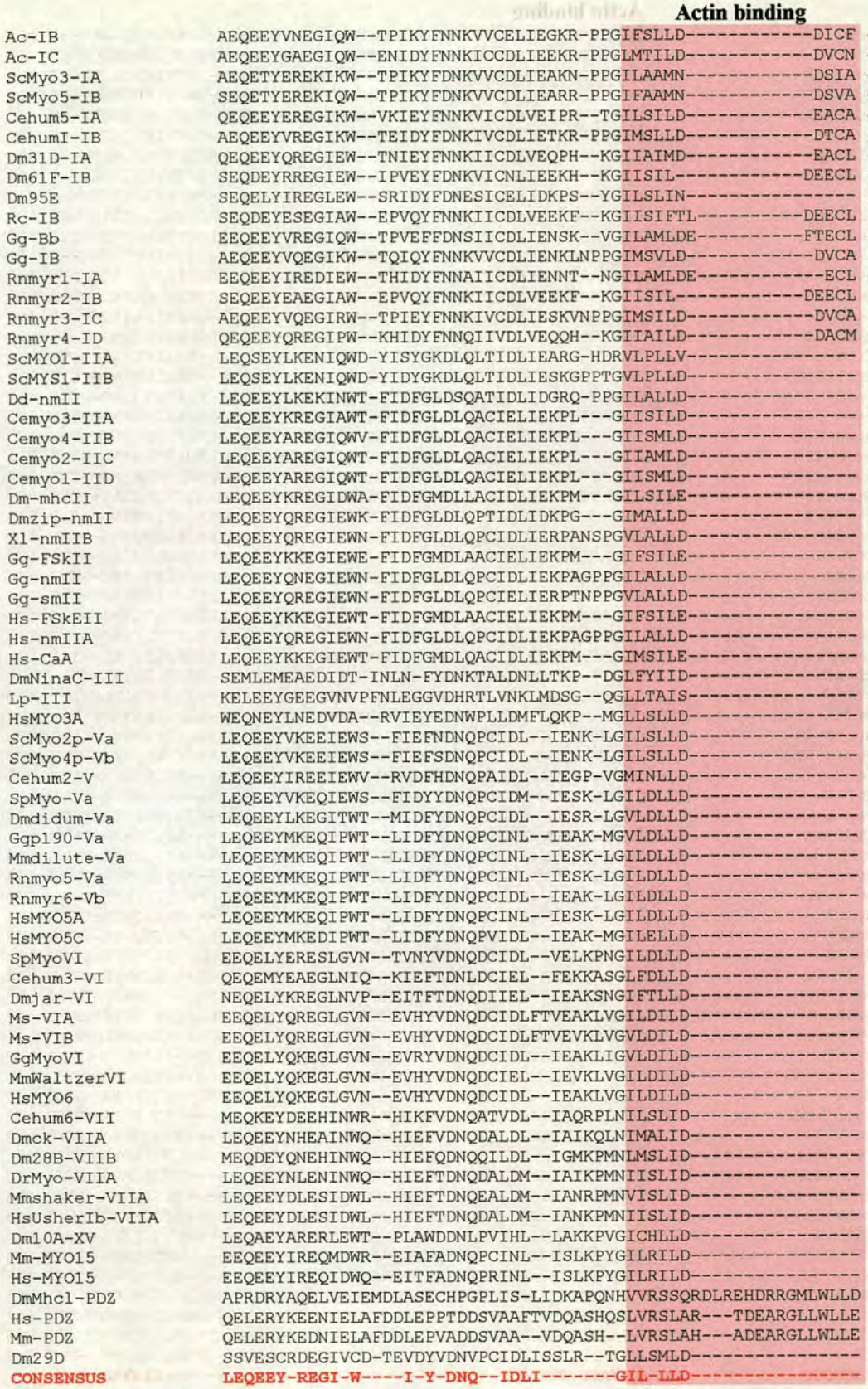
```
Ac-IB              MTAEEQRQVFH--LVAGILHLG--NVAFHDGGK---------GTAAVHD-------RTPF 283
Ac-IC              FTAEEQHEIFR--LVTAILYLG--NVQFVDDGK---------GGSTIAD-------SRPV 281
ScMyo3-IA          LVQEEQDQIFR--MLAAILWIG--NISFIENEE---------GNAQVGDTSVTDFVAYLL 294
ScMyo5-IB          LGQEEQDQIFR--MLAAILWIG--NVSFIENEE---------GNAQVRDTSVTDFVAYLL 294
Cehum5-IA          FDKQDVESMWS--VIAGLIHLG--NVRFIDGENS-------SGAVHIAEKAALQNAARCL 299
CehumI-IB          VNDQDQLEVLR--IVATVLHIG--NITFTEENN---------FAAVSGKD-YLEYPAFLL 291
Dm31D-IA           FSTDEVQTIWR--TIAAVLHLG--NVEFQTIED---------ELVISNKQHLKSTAKLL 292
Dm61F-IB           FGKEEQREIFG--IVASILHLG--NVGFTEVE---------G-NAKVNSRDLVVTAARLL 294
Dm95E              LSCEESNSIFR--VIAVVLKLG--NFIFVPITNID---G--TEGCQVSNVYEVQETAQLL 284
Rc-IB              FNDDDIEELLS--IVASVLHLGNFTVQFATDE---------HGHAQVTTENQIKYLARLL 305
Gg-Bb              FAPAEVTALLE--VTAVVLKLG--NVKFTLSSCFQ---ASGMEASSITEPRELQEISQLI 307
Gg-IB              IRGEDQQLVLQ--IVAGILHLG--NISFREEGN---------YARVENAD-SLAFPAYLL 292
Rnmyr1-IA          FSDPEAESVLE--VVAAVLKLG--NIEFKPES--R---MNGLDESKIKDKNELKEICELT 297
Rnmyr2-IB          FTEDEVEDLLS--IVASVLHLG--NIHFAADE---------DSNAQVTTENQLKYLTRLL 295
Rnmyr3-IC          IFSEEQTLVLQ--IVAGILHLG--NINFKEVGN---------YAAVESEE-FLAFPAYLL 292
Rnmyr4-ID          FKPEEIQTVYK--ILAAILHLG--NLKFIVDGD----------TPLIENGKVVSVIAELL 294
ScMYO1-IIA         SKDQI-RWIFQ--VVAIILLIG--NIEFVSDRAE--------QASFKN---DVSAICSNL 314
ScMYS1-IIB         SKDQI-RWIFQ--VVAIILLIG--NIEFVSDRAE--------QASFKN---DVSAICSNL 315
Dd-nmII            SQEEQ-MSIFK--IIAGILHLG--NIKFEKGAG---------EGAVLKDKTALNAASTVF 295
Cemyo3-IIA         EDNET-MDLYR--STAGIMHMG--EMKFKQRPRE--------EQAEPDGEEDALNAAAML 304
Cemyo4-IIB         SAVEK-QDCYR--LMSAHMHMG--NMKFKQRPRE--------EQAEPDGTVEAEKASNMY 304
Cemyo2-IIC         TPTEK-MECYR--LVAAMMHMG--NMKFKQRPRE--------EQAEPDGTDDAERAAKCF 310
Cemyo1-IID         SPTEK-MDCYR--LMSAHMHMG--NMKFKQRPRE--------EQAEPDGQVEAERACNMY 301
Dm-mhcII           TKQEK-EDVYR--ITAAVMHMG--GMKFKQRGRE--------EQAEQDGEEEGGRVSKLF 300
Dmzip-nmII         TSEDF-NSIFR--IVSAVLLFG--SMKFRQERNN--------DQATLPDNTVAQKIAHLL 344
Xl-nmIIB           SHEEI-LSMLK--VVSSVLQFG--NIVFKKERNT--------DQASMPENTAAQKLCHLL 315
Gg-FSkII           SADEK-TAIYK--LTGAVMHYG--NLKFKQKQRE--------EQAEPDGTEVADKAAYLM 303
Gg-nmII            PDEEQ-IGLLK--VISGVLQLG--NIVFKKERNT--------DQASMPDNTAAQKVSHLL 296
Gg-smII            TEEEQ-TSILR--VVSSVLQLG--NIVFKKERNT--------DQASMPDNTAAQKVCHLM 306
Hs-FSkEII          TPEEK-SGLYK--LTGAVMHYG--NMKFKQKQRE--------EQAEPDGTEVADKTAYLM 302
Hs-nmIIA           PEEEQ-MGLLR--VISGVLQLG--NIVFKKERNT--------DQASMPDNTAAQKVSHLL 296
Hs-CaA             TSEEK-AGVYK--LTGAIMHYG--NMKFKQKQRE--------EQAEPDGTEDADKSAYLM 303
DmNinaC-III        FNHKQLETVRK--VLAAILNIG--NIRFRQNGKYA---E--VENTDIV-----SRIAELL 298
Lp-III             FTEEEIFAIQK--VLAAILLLG--ETEIQN------------------------------ 268
HsMYO3A            FTMEQLGSIYS--ILAAILNVG--NIEFSSVATEH---Q--IDKSHISNHTALENCVSLL 300
ScMyo2p-Va         TKETQ-HQIFK--ILAALLHIG--NIEIKKT---------RNDASLSADEPNLKLACELL 301
ScMyo4p-Vb         NHETQ-LGIFK--ILAGLLHIG--NIEMKMT---------RNDASLSSEEQNLQIACELL 299
Cehum2-V           DEKQM-SDVFR--LLAGLLLLG--NVHFENG-------EGSSAVSASSCQEISRLCREFW 303
SpMyo-Va           KDDMM-KQIFR--IIGAVLHFG--NVEIQP--------DQHESCKIEVESESLPILCSLL 294
Dmdidum-Va         SIQQI-ADIVK--ILAGILHLG--NIQVSKKFNEGSEEEDSDSCDIFHNDIHLQITADLL 303
Ggp190-Va          SDSYQ-MGIFR--ILAGILHLG--NVEFAS--------RDSDSCAIPPKHDPLTIFCDLM 294
Mmdilute-Va        SESYQ-MGIFR--ILAGILHLG--NVGFAS--------RDSDSCTIPPKHEPLTIFCDLM 294
Rnmyo5-Va          SESYQ-MGIFR--ILAGILHLG--NVGFAS--------RDSDSCTIPPKHEPLIIFCDLM 294
Rnmyr6-Vb          RESHQ-ISIFK--IIASILHLG--SVEIQAE-------RDGDSCSISPQDEHLSNFCRLL 295
HsMYO5A            SESHQ-MGIFR--ILAGILHLG--NVGFTS--------RDADSCTIPPKHEPLCIFCDLM 294
HsMYO5C            KEDFQ-MDVFK--ILAAILHLG--NVQITA--------VGNERSSVSEDDSHLKVFCELL 294
SpMyoVI            KPQEE-YDILR--IVACVLHLG--NIAFTGN-----DSQEGGCGVSPDAQDSLKFVADLI 322
Cehum3-VI          SDDEK-MFIWS--TVAGILHLG--NIEFEENA----SDSRGGCMITSGTENSLMAAAELL 320
Dmjar-VI           SDTEK-LGIYS--LVAAVLHLG--NIAFEEIP----DDVRGGCQVSEASEQSLTITSGLL 321
Ms-VIA             DDTEK-LDLFRGFTVAGVLHLG--NIDFEEA-----GSTSGGCAIKNQSIQTLEHCAELL 331
Ms-VIB             DDTEK-LDLFRVFTVAGVLHLG--NIDFEET-----GSSSGGCVLKNQSGQTLEYCADLL 331
GgMyoVI            DDAEK-LDLFR--VVAGVLHLG--NIDFEEA-----GSTSGGCTLKAQSQPALECCAALL 321
MmWaltzerVI        GMMKKNFDLFR--VVAGVLHLG--NIDLEEA-----GSTSGGCNLKNKSAPSLEYCAELL 322
HsMYO6             DDEEK-LDLFR--VVAGVLHLG--NIDFEEA-----GSTSGGCNLKNKSAQSLEYCAELL 321
Cehum6-VII         NEQEI-GSIFK--LLASLLHIG--NIRFRQNTN------DNMESVDVADPSTLVRIAKLL 293
Dmck-VIIA          SDQEI-WEIIK--LLAALLHCG--NIKYKATVV------DNLDATEIPEHINVERVAGLL 293
Dm28B-VIIB         KPEEV-WSILS--LLAAILHLG--NLRFTATEV------ANLATAEIDDTPNLQRVAQLL 295
DrMyo-VIIA         TDKEN-WEISK--LLAAILHMG--NLRYEARTY------DNLDACEVVRCSALTTAAVLL 293
Mmshaker-VIIA      TDTEN-WEISK--LLAAILHMG--NLQYEARTF------ENLDACEVLFSPSLATAASLL 293
HsUsherIb-VIIA     TDTEN-WEISK--LLAAILHLG--NLQYEARTF------ENLDACEVLFSPSLATAASLL 293
Dm10A-XV           SEGER-EGIVR--VLAAVLHLG--NVYFHRRQLR-----HGQEGVEVGSDAEIKWAAHLL 292
Mm-MYO15           TSEDQ-DSIFR--ILASILHLG--NVYFEKHET------DAQEVASVVSAREIQAVAELL 293
Hs-MYO15           SSEDQ-DSIFR--ILASILHLG--NVYFEKYET------DAQEVASVVSAREIQAVAELL 293
DmMhcl-PDZ         DAKAV-RGIWS--ILAAIYHLG--IAGVTKLG------TGSTARTQFANPTAARKASGLL 297
Hs-PDZ             SPDEQ-KACWF--ILAAIYHLG--AAGATKEFT-----AAEAGRKQFARHEWAQKAAYLL 252
Mm-PDZ             SPEEQ-KTCWL--ILASIYHLG--AAGATKE-------AAEAGRKQFARHEWAQKAAYLL 296
Dm29D              ---IPFLDVVR--VLAAVLLLG--NVQFIDGGG---------LEVDVKGETELNSVASLL 283
CONSENSUS          ---E----IFR--ILAAILHLG--NI-F----------------------L---A-LL
```

```
Ac-IB              ALKNALLFRVLNTGGAG---AKKMS-TYNVPQNVEQA---------AS--ARDALAKTIY 328
Ac-IC              AVETALLYRTITTGEQG---RGRSS-VYSCPQDPLGA---------IY--SRDALSKALY 326
ScMyo3-IA          QVDASLLVKCLVERIMQTSHGMKRGSVYHVPLNPVQA---------TA--VRDALAKAIY 343
ScMyo5-IB          QIDSQLLIKSLVERIMETNHGMKRGSVYHVPLNIVQA---------DA--VRDALAKAIY 343
Cehum5-IA          NVTPDELAKSLSSQVVAAHGDIVKK-----QHDVNAA---------YY--TRDALAKALY 343
CehumI-IB          GLTSADIEAKLTGRKMESKWGTQKE-EIDMKLNVEQA---------SY--TRDAWVKAIY 339
Dm31D-IA           QVTETELSTALTKRVIAAGGNVMQK-----DHNATQA---------EY--GKDALAKAIY 336
Dm61F-IB           GVNASELEAALTHRTIDARGDVVTS-----PLNQELA---------IY--ARDALAKAVY 338
Dm95E              NMEAQILINCLTRANSTNSAQEDVG----CEMDARQA---------AT--NRNTLCRTLY 329
Rc-IB              SVDSTVLRESLIHKKIIAKGEELNS-----PLNLEQA---------AFTYARDALAKAIY 351
Gg-Bb              GLDPSTLEQALCSRTVKVRDESVLT-----ALSVSQF---------TGYYGRDALAKNIY 353
Gg-IB              GIDRDRLNDKITSRKMDSKWGGRSE-SITVTLNVEQA---------AY--TRDALAKGLY 340
Rnmyr1-IA          SIDQVVLERAFSFRTVEAKQEKVST-----TLNVAQ----------AYYARDALAKNLY 341
Rnmyr2-IB          GVEGTTLREALTHRKIIAKGEELLS-----PLNLEQA---------AY--ARDALAKAVY 339
Rnmyr3-IC          GINQDRLKEKLTSRQMDSKWGGKSE-SIHVTLNVEQA---------CY--TRDALAKALH 340
Rnmyr4-ID          STKADMVEKALLYRTVATGRDIIDK-----QHTEQEA---------SY--GRDAFAKAIY 338
ScMYO1-IIA         GVDEKDFQTAILRPRSKAGK-------EWVSQSKNST---------KLSSLLNALSRNLY 358
ScMYS1-IIB         GVDEKDFQTAILRPRSKAGK-------EWVSQSKNSQ---------QAKFILNALSRNLY 359
Dd-nmII            GVNPSVLEKALMEPRILAGR-------DLVAQHLNVE---------KSSSSRDALVKALY 339
Cemyo3-IIA         GIQAEEFLKALTKPRVRVGT-------EWVNKGQNLE---------QVNWAVSGLAKAIY 348
Cemyo4-IIB         GIGCEEFLKALTKPRVKVGT-------EWVSKGQNCE---------QVNWAVGAMAKGLY 348
Cemyo2-IIC         GIDSEEFLKALTRPRVKVGN-------EWVNKGQKIE---------QVNWAVGAMAKGLY 354
Cemyo1-IID         GIDVVQFLKALVSPRVKVGT-------EWVSKGQNVD---------QVHWAIGAMAKGLY 345
Dm-mhcII           GCDTAELYKNLLKPRIKVGN-------EFVTQGRNVQ---------QVTNSIGALCKGVF 344
Dmzip-nmII         GLSVTDMTRAFLTPRIKVGR-------DFVTKAQTKE---------QVEFAVEAIAKACY 388
Xl-nmIIB           GLNIMEFTRAILTPRIKVGR-------DYVQKAQTKE---------QADFAVEALRKATY 359
Gg-FSkII           GLNSAELLKALCYPRVKVGN-------EFVTKGQTVS---------QVHNSVGALAKAVY 347
Gg-nmII            GINVTDFTRGILTPRIKVGR-------DYVQKAQTKE---------QADFAIEALAKATY 340
Gg-smII            GINVTDFTRSILTPRIKVGR-------DVVQKAQTKE---------QADFAIEALAKAKF 350
Hs-FSkEII          GLNSSDLLKALCFPRVKVGN-------EYVTKGQTVD---------QVHHAVNALSKSVY 346
Hs-nmIIA           GINVTDFTRGILTPRIKVGR-------DYVQKAQTKE---------QADFAIEALAKATY 340
Hs-CaA             GLNSADLLKGLCHPRVKVGN-------EYVTKGQSVQ---------QVYYSIGALAKAVY 347
DmNinaC-III        RVDEKKFMWSLTNFIMVKGGIAERR-----QYTTEEA---------RD--ARDAVASTLY 342
Lp-III             -----------SAAFKLLGAESSEL-----ENTLTQD---------VN--ARDVYARAMY 301
HsMYO3A            CIRADELQEALTSHCVVTRGETIIR-----PNTVEKA---------TD--VRDAMAKTLY 344
ScMyo2p-Va         GIDAYNFAKWVTKKQIITRS-------EKIVSNLNYS---------QALVAKDSVAKFIY 345
ScMyo4p-Vb         GIDPFNFAKWIVKKQIVTRS-------EKIVTNLNYN---------QALIARDSVAKFIY 343
Cehum2-V           KISESDLRIWLTRREIRAVN-------EIVTKPLTKN---------EAVRSRDALTKMLY 347
SpMyo-Va           GIEEDAMRKWLCNRKIVTVQ-------EVLTKPLRRD---------MAVFSRDALAKHIY 338
Dmdidum-Va         RVSADDLRRWLLMRKIESVN-------EYVLIPNSIE---------AAQAARDALAKHIY 347
Ggp190-Va          GVDYEEMAHWLCHRKLATAT-------ETYIKPISKL---------HAINARDALAKHIY 338
Mmdilute-Va        GVDYEEMCHWLCHRKLATAT-------ETYIKPISKL---------QATNARDALAKHIY 338
Rnmyo5-Va          GVDYEEMCHWLCHRKLATAT-------ETYIKPISKL---------QATNARDALAKHIY 338
Rnmyr6-Vb          GIEHSQMEHWLCHRKLVTTS-------ETYVKTMSLQ---------QVVNARNALAKHIY 339
HsMYO5A            GVDYEEMCHWLCHRKLATAT-------ETYIKPISKL---------QATNARDALAKHIY 338
HsMYO5C            GLESGRVAQWLCNRKIVTSS-------ETVVKPMTRP---------QAVNARDALAKKIY 338
SpMyoVI            KVDAKELEQSLTSRTMQARG-------TVIKVPLKKE---------QASNARDALAKAVY 366
Cehum3-VI          GLEPEEMKLGLCARIMQTTK--GGVKGTLIRVPLKAH---------EASAGRDALAKAIY 369
Dmjar-VI           GVDQTELRTALVSRVMQSKG--GGFKGTVIMVPLKIY---------EASNARDALAKAIY 370
Ms-VIA             GLEEGDLRVSLTTRVMLTTAFTGGAKGTLIKVPLKVE---------QANNSRDALAKAVY 382
Ms-VIB             GLDQDDLRVSLTTRIMLTTAFTGGAKGTVIKVPLKVE---------QANNARDALAKAVY 382
GgMyoVI            GLDEEDLRVSLTTRVMLTTA--GGAKGTVIKVPLKVE---------QANNARDALAKTVY 370
MmWaltzerVI        GLDQDDLRVSLTTRVMLTTA--GGTKGTVIKVPLKVE---------QANNARDALAKTVY 371
HsMYO6             GLDQDDLRVSLTTRVMLTTA--GGTKGTVIKVPLKVE---------QANNARDALAKTVY 370
Cehum6-VII         QLHEQNLLDAITTKSLVTRE-------ERVISRLNGQ---------QAVDARDALAKAIY 337
Dmck-VIIA          GLPIQPLIDALTRRTLFAHG-------ETVVSTLSRD---------QSVDVRDAFVKGIY 337
Dm28B-VIIB         GIPISALNAALTQRTIFVHG-------EHVTTSLSKE---------AAIEGRDAFVKSLY 339
DrMyo-VIIA         EVDLKDLMNCLTSRTIITRG-------ETVSTPLSTE---------QALDVRDAFVKGIY 337
Mmshaker-VIIA      EVNPPDLMSCLTSRTLITRG-------ETVSTPLSRE---------QALDVRDAFVKGIY 337
HsUsherIb-VIIA     EVNPPDLMSCLTSRTLITRG-------ETVSTPLSRE---------QALDVRDAFVKGIY 337
Dm10A-XV           HISADGLHRALTSRTTEARA-------ERLHTPLGID---------QALDARDAFAKALY 336
Mm-MYO15           QVSPEGLQKAITFKVTETIR-------EKIFTPLTVE---------SAVDARDAIAKVLY 337
Hs-MYO15           QISPEGLQKAITFKVTETMR-------EKIFTPLTVE---------SAVDARDAIAKVLY 337
DmMhc1-PDZ         GVNLEDLSSAAFGLTQPN-APNGGLSPSKSP-TSDTG-------HEWAWECLEALVIGLY 348
Hs-PDZ             GCSLEELSSAIFKHQHKGGTLQRSTSFRQGPEESGLGFTDGTGPKLSALECLEGMAAGLY 312
Mm-PDZ             GCSLEELSSAIFKHQLKGGTLQRSTSFRQGPEESGLG----EGTKLSALECLEGMASGLY 352
Dm29D              GVPPAALFRGLTTRTHNVRG--------QLVKSVCGDG--------DANMTRDCLAKALY 327
CONSENSUS          GV---DL---LT-R-I----------E-V---L----------A--ARDALAKAIY
```


Drosophila Myo29D insert
Ac-IB SRMFDWIVSKVNEALQKQGGSGD---------- 351
Ac-IC SRMFDYIIQRVNDAMYID----D---------- 345
ScMyo3-IA NNLFDWIVDRVNVSLQAFP---------- 362
ScMyo5-IB NNLFDWIVSRVNKSLQAFP---------- 362
Cehum5-IA ERLFSWMVSKVNEAISVQNSSRYS---------- 367
CehumI-IB ARLFDYLVKKVNDAMNITS-QST---------- 361
Dm31D-IA DRLFTWIISRINRAILFRGSKTQA---------- 360
Dm61F-IB DRLFSWLVQRLNISLQAKETRAS---------- 361
Dm95E SRLFTWLVNKINESLKSTQREKN---------- 352
Rc-IB GRTFSWLVSKINKSLAYKGTDMHK---------- 375
Gg-Bb SRLFDWLVNRINTSIQVKPGKQR---------- 376
Gg-IB ARVFDFLVESINRAM-----QKP---------- 358
Rnmyr1-IA SRLFSWLVNRINESIKAQTKVRK---------- 364
Rnmyr2-IB SRTFTWLVRKINRSLASKDAESPS---------- 363
Rnmyr3-IC ARVFDFLVDSINKAM-----EKD---------- 358
Rnmyr4-ID ERLFCWIVTRINDIIEVKNYDTTV---------- 362
ScMYO1-IIA ERLFGYIVDMINKNLDHG---------- 376
ScMYS1-IIB ERLFGYIVDMINKNLDHG---------- 377
Dd-nmII GRLFLWLVKKINNVLCQE---------- 357
Cemyo3-IIA ARMFKWIITRCNKTLDAKE---------- 367
Cemyo4-IIB SRVFNWLVKKCNLTLDQKG---------- 367
Cemyo2-IIC SRIFNWLVKKCNQTLDQKG---------- 373
Cemyo1-IID ARVFHWLVKKCNLTLDQKG---------- 364
Dm-mhcII DRLFKWLVKKCNETLDTQ---------- 362
Dmzip-nmII ERMFKWLVNRINRSLDRTK---------- 407
Xl-nmIIB ERLFRWLVYRINKALDRTK---------- 378
Gg-FSkII EKMFLWMVIRINQQLDTK---------- 365
Gg-nmII EQMFRWLVMRINKALDKTK---------- 359
Gg-smII ERLFRWILTRVNKALDKTK---------- 369
Hs-FSkEII EKLFLWMVTRINQQLDTK---------- 364
Hs-nmIIA ERMFRWLVLRINKALDKTK---------- 359
Hs-CaA EKMFNWMVTRINATLETK---------- 365
DmNinaC-III SRLVDFIINRINMNMSFPRAVFG---------- 365
Lp-III LRLFSWIVAVVNRQLSFSRLVFG---------- 324
HsMYO3A GRLFSWIVNCINSLLKHDSSPSGN---------- 368
ScMyo2p-Va SALFDWLVENINTVLCNP----AVN---------- 366
ScMyo4p-Vb STLFDWLVDNINKTLYDP----ELD---------- 364
Cehum2-V SHLFGWLVDKINEALNEKDKLDGTN---------- 372
SpMyo-Va SQLFTWIVEQINKAMD---------- 354
Dmdidum-Va AKLFQYIVGVLNKSLN---------- 363
Ggp190-Va ANLFNWIVDHVNKALH---------- 354
Mmdilute-Va AKLFNWIVDHVNQALH---------- 354
Rnmyo5-Va AKLFNWIVGHVNQALH---------- 354
Rnmyr6-Vb AQLFSWIVEHINKALQ---------- 355
HsMYO5A AKLFNWIVDNVNQALH---------- 354
HsMYO5C AHLFDFIVERINQALQ---------- 354
SpMyoVI SHLFDHIVSRINECFP---------- 382
Cehum3-VI SKLFDWLVAQINKSIP---------- 385
Dmjar-VI SRLFDRIVGLINQSIP---------- 386
Ms-VIA SRLFDHVVTRVNQCFP---------- 398
Ms-VIB SRLFDHVVKRVNQCFP---------- 398
GgMyoVI SHLFDHVVNRVNQCFP---------- 386
MmWaltzerVI SHLFDHVVNRVNQCFP---------- 387
HsMYO6 SHLFDHVVNRVNQCFP---------- 386
Cehum6-VII GKLFIHIVRRVNDAIYKPSQ---------- 357
Dmck-VIIA GRMFVHIVRKINTAIFKPRG---------- 357
Dm28B-VIIB DGIFVRIVRRINETINKQVD---------- 359
DrMyo-VIIA GRLFVWIVEKINAAIYKPPSL---------- 358
Mmshaker-VIIA GRLFVWIVEKINAAIYKPPPL---------- 358
HsUsherIb-VIIA GRLFVWIVDKINAAIYKPPSQ---------- 358
Dm10A-XV AGLFNWLVSRIN-SIVQKGG---------- 355
Mm-MYO15 ALLFGWLITRVN-ALVSP---------- 354
Hs-MYO15 ALLFSWLITRVN-ALVSP---------- 354
DmMhcl-PDZ SEALAAVVALINRQICTSS---------- 367
Hs-PDZ SELFTLLVSLVNRALKSSQ---------- 331
Mm-PDZ SELFTLLISLVNRALKSSQ---------- 371
Dm29D CRTVATIVRRANSLKRLGSTLGTLSSDSNESVHNQADVASQHASTIGGGNAGSKSMAALN 387
CONSENSUS -RLF-WIV-RIN--L----------

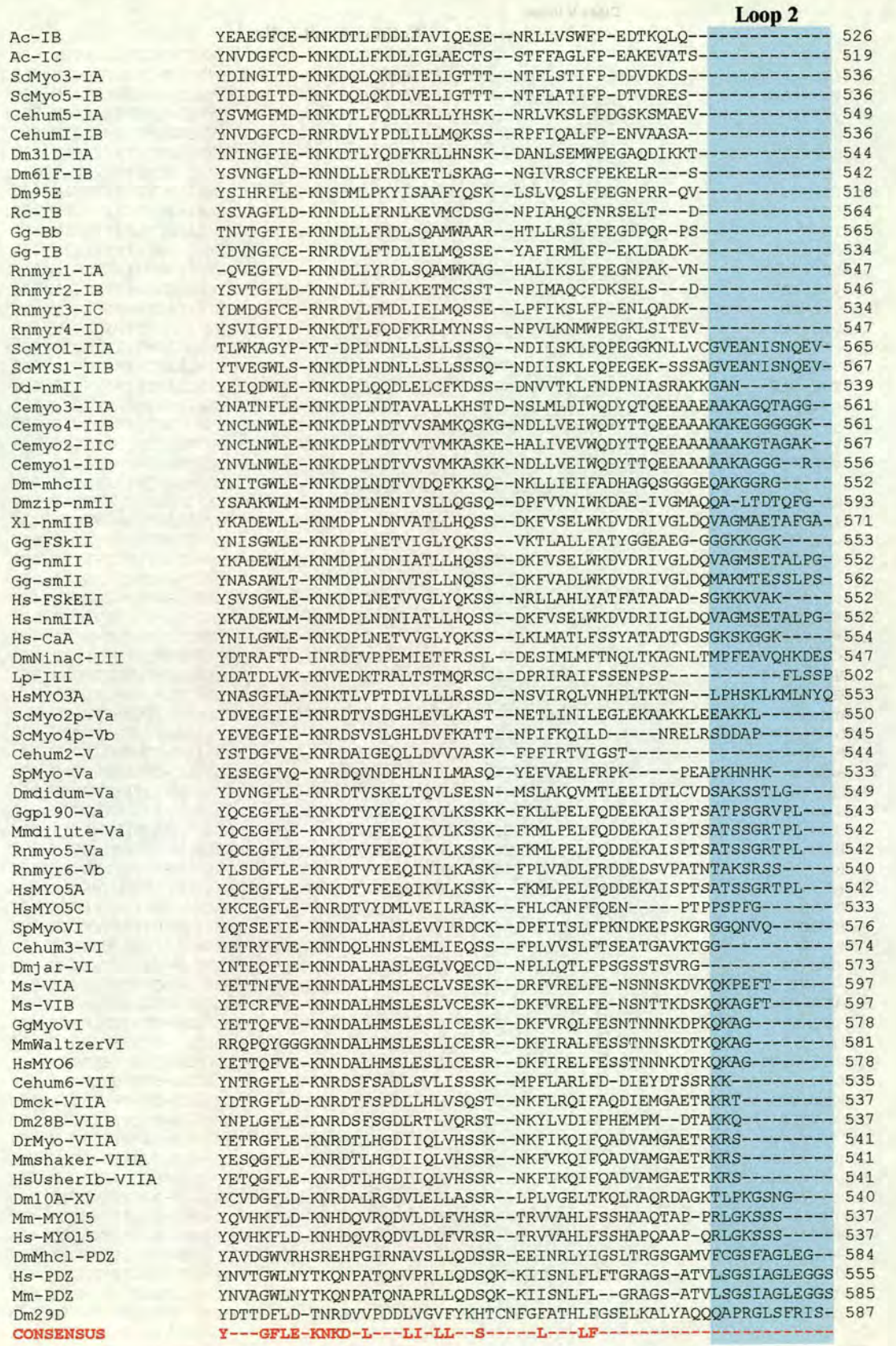
Switch 2

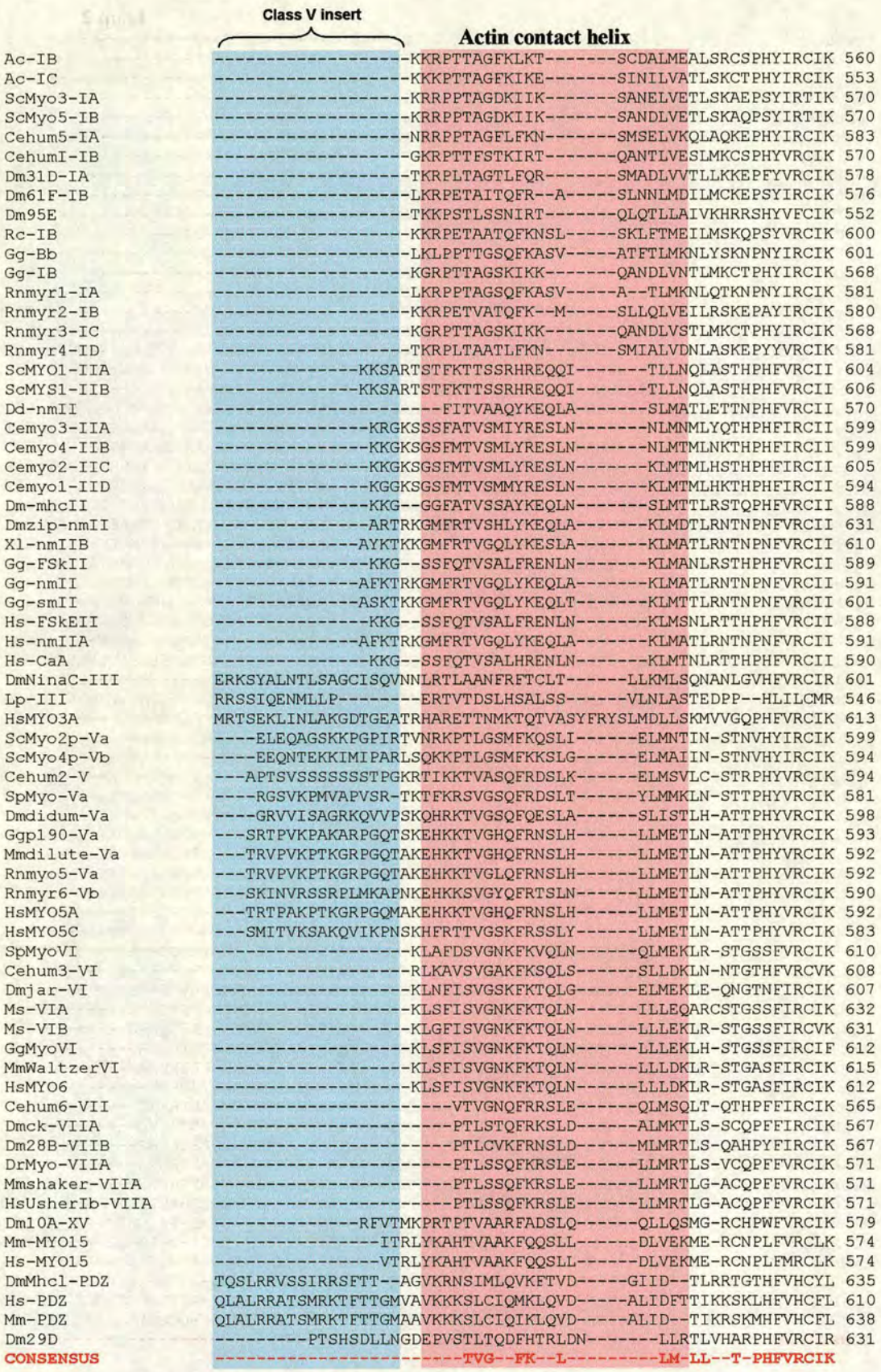
```
Ac-IB                ------HNNNMIGVLDIF--GFEIFE------QNGFE--QFCINYVNEKLQQYFIELTLK 395
Ac-IC                ------PEALTTGILDIY--GFEIFG------KNGFE--QLCINFVNEKLQQIFIQLTLK 389
ScMyo3-IA            ------GADKSIGILDIY--GFEIFE------HNSFE--QICINYVNEKLQQIFIQLTLK 406
ScMyo5-IB            ------GAEKSIGILDIY--GFEIFE------HNSFE--QICINYVNEKLQQIFIQLTLK 406
Cehum5-IA            -----K--SHVIGVLDIY--GFEIFG------TNSFE--QLCINYCNEKLQQLFIELVLK 410
CehumI-IB            ------SDNFSVGILDIY--GFEIFN------NNGFE--QFCINFVNEKLQQIFIELTLK 405
Dm31D-IA             -----R-FNSVIGVLDIY--GFEIFD------SNSFE--QFCINYCNEKLQQLFIELVLK 404
Dm61F-IB             -------RNNVMGILDIY--GFEIFQ------KNSFE--QFCINFCNEKLQQLFIELTLK 404
Dm95E                -----------LALLDFY--GFEALD------HNSFE--QFAINYSAEKIHQNFVFHVLR 391
Rc-IB                -----LGSASVIGLLDIY--GFEVFQ------HNSFEFTQFCINFCNEKLQQLFIELTLK 422
Gg-Bb                ---------KVMGVLDIY--GFEIFQ------DNGFEQFIFTINYCNEKLQQIFILMTLK 419
Gg-IB                ------YEEYSIGVLDIY--GFEIFQ------KNGFE--QFCINFVNEKLQQIFIELTLK 402
Rnmyr1-IA            ---------KVMGVLDIY--GFEIFE------DNSFEQFI--INYCNEKLQQIFIELTLK 405
Rnmyr2-IB            -----WRSTTVLGLLDIY--GFEVFQ------HNSFE--QFCINYCNEKLQQLFIELTLK 408
Rnmyr3-IC            ------HEEYNIGVLDIY--GFEIFQ------KNGFE--QFCINFVNEKLQQIFIELTLK 402
Rnmyr4-ID            -----HGKNTVIGVLDIY--GFEIFD------NNSFE--QFCINYCNEKLQQLFIQLVLK 407
ScMYO1-IIA           -----SATLNYIGLLDIA--GFEIFE------NNSFE--QLCINYTNEKLQQFFNNHMFV 421
ScMYS1-IIB           -----SATLNYIGLLDIA--GFEIFE------NNSFE--QLCINYTNEKLQQFFNNHMFV 422
Dd-nmII              ------RKAYFIGVLDIS--GFEIFK------VNSFE--QLCINYTNEKLQQFFNHHMFK 401
Cemyo3-IIA           -----IERKHFIGVLDIA--GFEIFD------LNSFE--QLWINFVNERLQQFFNHHMFV 412
Cemyo4-IIB           -----IDRDYFIGVLDIA--GFEIFD------FNSFE--QLWINFVNEKLQQFFNHHMFV 412
Cemyo2-IIC           -----ISRDHFIGVLDIA--GFEIFD------FNSFE--QLWINFVNEKLQQFFNHHMFV 418
Cemyo1-IID           -----IDRDYFIGVLDIA--GFEIFD------FNSFE--QLWINFVNEKLQQFFNHHMFV 409
Dm-mhcII             -----QKRQHFIGVLDIA--GFEIFE------YNGFE--QLCINFTNEKLQQFFNHHMFV 407
Dmzip-nmII           -----RQGASFIGILDMA--GFEIFE------LNSFE--QLCINYTNEKLQQLFNHTMFI 452
Xl-nmIIB             -----RQGASFIGILDIA--GFEIFE------LNSFE--QLCINYTNEKLQQLFNHTMFI 423
Gg-FSkII             -----QPRQYFIGVLDIA--GFEIFD------FNSFE--QLCINFTNEKLQQFFNHHMFV 410
Gg-nmII              -----RQGASFIGILDIA--GFEIFE------LNSFE--QLCINYTNEKLQQLFNHTMFI 404
Gg-smII              -----RQGASFLGILDIA--GFEIFE------INSFE--QLCINYTNEKLQQLFNHTMFI 414
Hs-FSkEII            -----LPRQHFIGVLDIA--GFEIFE------YNSLE--QLCINFTNEKLQQFFNHHMFV 409
Hs-nmIIA             -----RQGASFIGILDIA--GFEIFD------LNSFE--QLCINYTNEKLQQLFNHTMFI 404
Hs-CaA               -----QPRQYFIGVLDIA--GFEIFD------FNSFE--QLCINFTNEKLQQFFNHHMFV 410
DmNinaC-III          -------DTNAIIIHDMF--GFECFN------RNGLE--QLMINTLNEQMQYHYNQRIFI 408
Lp-III               -------DVYSVTVIDSP--GFENGL------HNSLH--QLCANVISDNLQNYIQQIIFF 367
HsMYO3A              ------GDELSIGILDIF--GFENFK------KNSFE--QLCINIANEQIQYYYNQHVFA 412
ScMyo2p-Va           -----DQISSFIGVLDIY--GFEHFE------KNSFE--QFCINYANEKLQQEFNQHVFK 411
ScMyo4p-Vb           ---QQDHVFSFIGILDIY--GFEHFE------KNSFE--QFCINYANEKLQQEFNQHVFK 411
Cehum2-V             ---QKKRPDRFIGVLDIY--GFETFD------VNSFE--QFSINYANEKLQQQFNQHVFK 419
SpMyo-Va             ---TQAKTTNFIGVLDIY--GFEMFE------INSFE--QFCINYANEKLQQQFTQHVFK 401
Dmdidum-Va           ---NGSKQCSFIGVLDIY--GFETFE------VNSFE--QFCINYANEKLQQQFNQHVFK 410
Ggp190-Va            ---STVKQHSFIGVLDIY--GFETFE------INSFE--QFCINYANEKLQQQFNMHVFK 401
Mmdilute-Va          ---SAVKQHSFIGVLDIY--GFETFE------INSFE--QFCINYANEKLQQQFNMHVFK 401
Rnmyo5-Va            ---SAVKQHSFIGVLDIY--GFETFE------INSFE--QFCINYANEKLQQQFNMHVFK 401
Rnmyr6-Vb            ---TSLKQHSFIGVLDIY--GFETFE------INSFE--QFCINYANEKLQQQFNSHVFK 402
HsMYO5A              ---SAVKQHSFIGVLDIY--GFETFE------INSFE--QFCINYANEKLQQQFNMHVFK 401
HsMYO5C              ---FSGKQHTFIGVLDIY--GFETFD------VNSFE--QFCINYANEKLQQQFNMHVFK 401
SpMyoVI              ----FKTSANFIGLLDIA--GFEFFT------LNSYE--QFCINYCNEKLQQFFNLRVLK 428
Cehum3-VI            ----FEKSTGYIGVLDVA--GFEYFA------VNSFE--QFCINFCNEKLQHFFNERILK 431
Dmjar-VI             ----FQASNFYIGVLDIA--GFEYFT------VNSFE--QFCINYCNEKLQKFFNDNILK 432
Ms-VIA               ----FDSSANFIGVLDIAFTGFEYFE------HNSFE--QFCINYCNEKLQQFFNERILK 446
Ms-VIB               ----FETSSNFIGVLDIAFTGFEYFE------HNSFE--QFCINYCNEKLQQFFNERILK 446
GgMyoVI              ----FETSSFFIGVLDIA--GFEYFE------HNSFE--QFCINYCNEKLQQFFNERILK 432
MmWaltzerVI          ----FETSSYFIGVLDIA--GFEYFE------HNSFE--QFCINYCNEKLQQFFNERILK 433
HsMYO6               ----FETSSYFIGVLDIA--GFEYFE------HNSFE--QFCINYCNEKLQQFFNERILK 432
Cehum6-VII           ------SRRTSIGILDIF--GFENFE------SNSFE--QLCINFANETLQQFFVHHVFK 401
Dmck-VIIA            ------TSRNAIGVLDIF--GFENFD------QNSFE--QFCINYANENLQQFFVQHIFK 401
Dm28B-VIIB           ------QPMNSIGVLDIF--GFENFD------NNSFE--QLCINYANENLQQFFVGHIFK 403
DrMyo-VIIA           ---ELKAVRRSIGLLDIF--GFENFM------VNSFE--QLCINFANENLQQFFVRHVFK 405
Mmshaker-VIIA        ---EVKNSRRSIGLLDIF--GFENFT------VNSFE--QLCINFANEHLQQFFVRHVFK 405
HsUsherIb-VIIA       ---DVKNSRRSIGLLDIF--GFENFA------VNSFE--QLCINFANEHLQQFFVRHVFK 405
Dm10A-XV             -----THDAHRISILDIF--GFEDLA------ENSFE--QLCINYANENLQLYFNKHVFK 400
Mm-MYO15             -----KQDTLSIAILDIY--GFEDLS------FNSFE--QLCINYANENLQYLFNKIVFQ 399
Hs-MYO15             -----RQDTLSIAILDIY--GFEDLS------FNSFE--QLCINYANENLQYLFNKIVFQ 399
DmMhcl-PDZ           ------HTIASIMLIDTP--GFQNPA--SCGQQVGATLADLRHNYLQERLQMLFHHTTLV 417
Hs-PDZ               ------HSLCSMMIVDTP--GFQNPEQFTGGSARGASFEELCHNYTQDRLQRLFHERTFV 383
Mm-PDZ               ------HSLCSMMIVDTP--GFQNPE--WGGSARGASFEELCHNYAQDRLQRLFHERTFL 421
Dm29D                NAVRHATSDGFIGILDMF--GFEEPS-------PHAHLEHLCINLCAETMQHFYNTHIFK 438
CONSENSUS            ----------FIGVLDIY--GFE-FE-------NSFE--QLCINY-NEKLQQ-FN--VFK
```

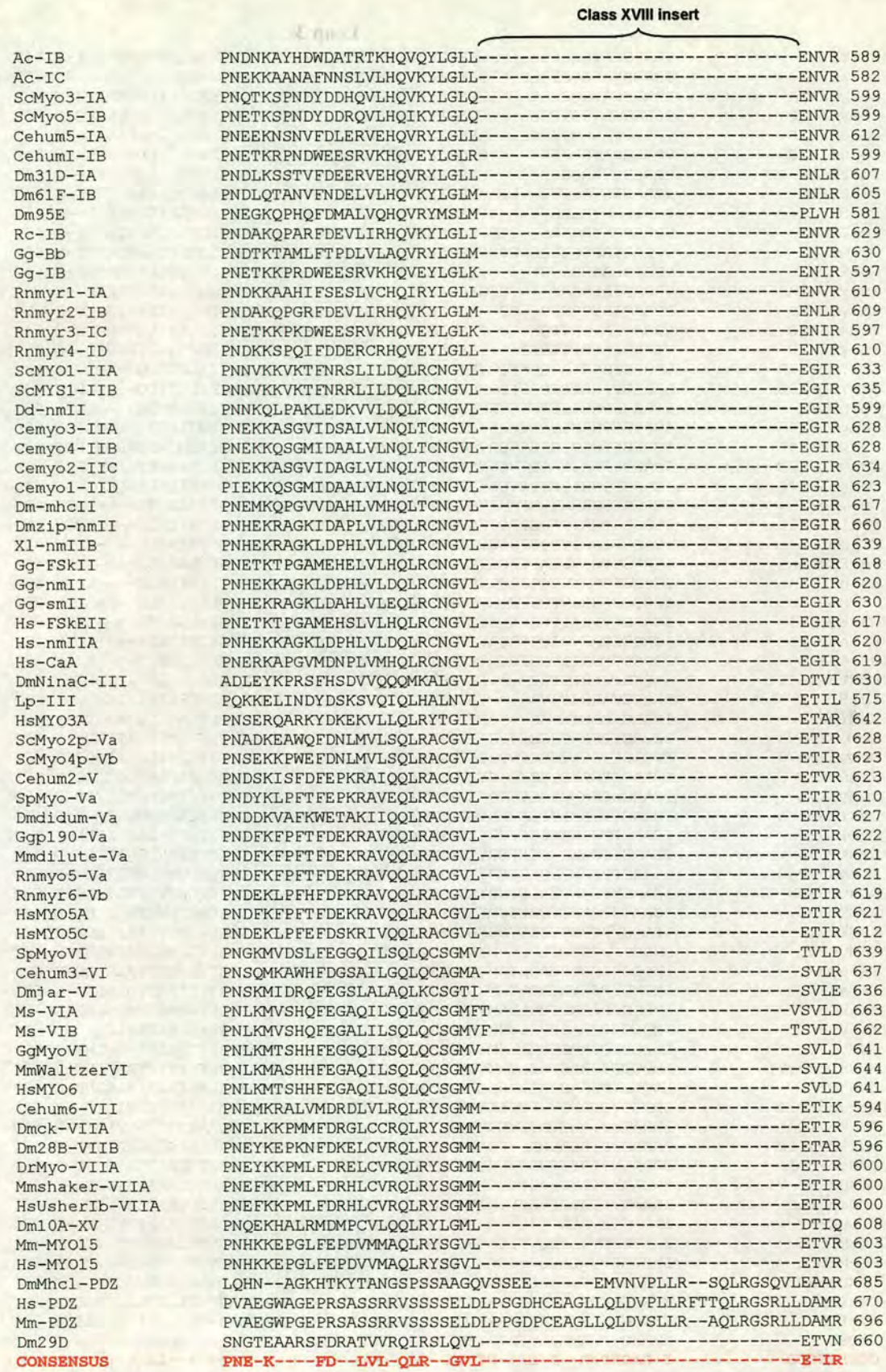

Actin binding
Ac-IB AEQEEYVNEGIQW--TPIKYFNNKVVCELIEGKR-PPGIFSLLD----------DICF 440
Ac-IC AEQEEYGAEGIQW--ENIDYFNNKICCDLIEEKR-PPGLMTILD----------DVCN 434
ScMyo3-IA AEQETYEREKIKW--TPIKYFDNKVVCDLIEAKN-PPGILAAMN----------DSIA 451
ScMyo5-IB SEQETYEREKIQW--TPIKYFDNKVVCDLIEARR-PPGIFAAMN----------DSVA 451
Cehum5-IA QEQEEYEREGIKW--VKIEYFNNKVICDLVEIPR--TGILSILD----------EACA 454
CehumI-IB AEQEEYVREGIKW--TEIDYFDNKIVCDLIETKR-PPGIMSLLD----------DTCA 450
Dm31D-IA QEQEEYQREGIEW--TNIEYFNNKIICDLVEQPH--KGIIAIMD----------EACL 448
Dm61F-IB SEQDEYRREGIEW--IPVEYFDNKVICNLIEEKH--KGIISIL----------DEECL 448
Dm95E SEQELYIREGLEW--SRIDYFDNESICELIDKPS--YGILSLIN---------- 431
Rc-IB SEQDEYESEGIAW--EPVQYFNNKIICDLVEEKF--KGIISIFTL----------DEECL 468
Gg-Bb EEQEEYVREGIQW--TPVEFFDNSIICDLIENSK--VGILAMLDE----------FTECL 465
Gg-IB AEQEEYVQEGIKW--TQIQYFNNKVVCDLIENKLNPPGIMSVLD----------DVCA 448
Rnmyr1-IA EEQEEYIREDIEW--THIDYFNNAIICDLIENNT--NGILAMLDE----------ECL 449
Rnmyr2-IB SEQEEYEAEGIAW--EPVQYFNNKIICDLVEEKF--KGIISIL----------DEECL 452
Rnmyr3-IC AEQEEYVQEGIRW--TPIEYFNNKIVCDLIESKVNPPGIMSILD----------DVCA 448
Rnmyr4-ID QEQEEYQREGIPW--KHIDYFNNQIIVDLVEQQH--KGIIAILD----------DACM 451
ScMYO1-IIA LEQSEYLKENIQWD-YISYGKDLQLTIDLIEARG-HDRVLPLLV---------- 463
ScMYS1-IIB LEQSEYLKENIQWD-YIDYGKDLQLTIDLIESKGPPTGVLPLLD---------- 465
Dd-nmII LEQEEYLKEKINWT-FIDFGLDSQATIDLIDGRQ-PPGILALLD---------- 443
Cemyo3-IIA LEQEEYKREGIAWT-FIDFGLDLQACIELIEKPL---GIISILD---------- 452
Cemyo4-IIB LEQEEYAREGIQWV-FIDFGLDLQACIELIEKPL---GIISMLD---------- 452
Cemyo2-IIC LEQEEYAREGIQWT-FIDFGLDLQACIELIEKPL---GIIAMLD---------- 458
Cemyo1-IID LEQEEYAREGIQWT-FIDFGLDLQACIELIEKPL---GIISMLD---------- 449
Dm-mhcII LEQEEYKREGIDWA-FIDFGMDLLACIDLIEKPM---GILSILE---------- 447
Dmzip-nmII LEQEEYQREGIEWK-FIDFGLDLQPTIDLIDKPG---GIMALLD---------- 492
Xl-nmIIB LEQEEYQREGIEWN-FIDFGLDLQPCIDLIERPANSPGVLALLD---------- 466
Gg-FSkII LEQEEYKKEGIEWE-FIDFGMDLAACIELIEKPM---GIFSILE---------- 450
Gg-nmII LEQEEYQNEGIEWN-FIDFGLDLQPCIDLIEKPAGPPGILALLD---------- 447
Gg-smII LEQEEYQREGIEWN-FIDFGLDLQPCIELIERPTNPPGVLALLD---------- 457
Hs-FSkEII LEQEEYKKEGIEWT-FIDFGMDLAACIELIEKPM---GIFSILE---------- 449
Hs-nmIIA LEQEEYQREGIEWN-FIDFGLDLQPCIDLIEKPAGPPGILALLD---------- 447
Hs-CaA LEQEEYKKEGIEWT-FIDFGMDLQACIDLIEKPM---GIMSILE---------- 450
DmNinaC-III SEMLEMEAEDIDT-INLN-FYDNKTALDNLLTKP--DGLFYIID---------- 448
Lp-III KELEEYGEEGVNVPFNLEGGVDHRTLVNKLMDSG--QGLLTAIS---------- 409
HsMYO3A WEQNEYLNEDVDA--RVIEYEDNWPLLDMFLQKP--MGLLSLLD---------- 452
ScMyo2p-Va LEQEEYVKEEIEWS--FIEFNDNQPCIDL--IENK-LGILSLLD---------- 450
ScMyo4p-Vb LEQEEYVKEEIEWS--FIEFSDNQPCIDL--IENK-LGILSLLD---------- 450
Cehum2-V LEQEEYIREEIEWV--RVDFHDNQPAIDL--IEGP-VGMINLLD---------- 458
SpMyo-Va LEQEEYVKEQIEWS--FIDYYDNQPCIDM--IESK-LGILDLLD---------- 440
Dmdidum-Va LEQEEYLKEGITWT--MIDFYDNQPCIDL--IESR-LGVLDLLD---------- 449
Ggp190-Va LEQEEYMKEQIPWT--LIDFYDNQPCINL--IEAK-MGVLDLLD---------- 440
Mmdilute-Va LEQEEYMKEQIPWT--LIDFYDNQPCINL--IESK-LGILDLLD---------- 440
Rnmyo5-Va LEQEEYMKEQIPWT--LIDFYDNQPCINL--IESK-LGILDLLD---------- 440
Rnmyr6-Vb LEQEEYMKEQIPWT--LIDFYDNQPCIDL--IEAK-LGILDLLD---------- 441
HsMYO5A LEQEEYMKEQIPWT--LIDFYDNQPCINL--IESK-LGILDLLD---------- 440
HsMYO5C LEQEEYMKEDIPWT--LIDFYDNQPVIDL--IEAK-MGILELLD---------- 440
SpMyoVI EEQELYERESLGVN--TVNYVDNQDCIDL--VELKPNGILDLLD---------- 468
Cehum3-VI QEQEMYEAEGLNIQ--KIEFTDNLDCIEL--FEKKASGLFDLLD---------- 471
Dmjar-VI NEQELYKREGLNVP--EITFTDNQDIIEL--IEAKSNGIFTLLD---------- 472
Ms-VIA EEQELYQREGLGVN--EVHYVDNQDCIDLFTVEAKLVGILDILD---------- 488
Ms-VIB EEQELYQREGLGVN--EVHYVDNQDCIDLFTVEVKLVGVLDILD---------- 488
GgMyoVI EEQELYQKEGLGVN--EVRYVDNQDCIDL--IEAKLIGVLDILD---------- 472
MmWaltzerVI EEQELYQKEGLGVN--EVHYVDNQDCIEL--IEVKLVGILDILD---------- 473
HsMYO6 EEQELYQKEGLGVN--EVHYVDNQDCIDL--IEAKLVGILDILD---------- 472
Cehum6-VII MEQKEYDEEHINWR--HIKFVDNQATVDL--IAQRPLNILSLID---------- 441
Dmck-VIIA LEQEEYNHEAINWQ--HIEFVDNQDALDL--IAIKQLNIMALID---------- 441
Dm28B-VIIB MEQDEYQNEHINWQ--HIEFQDNQQILDL--IGMKPMNLMSLID---------- 443
DrMyo-VIIA LEQEEYNLENINWQ--HIEFTDNQDALDM--IAIKPMNIISLID---------- 445
Mmshaker-VIIA LEQEEYDLESIDWL--HIEFTDNQEALDM--IANRPMNVISLID---------- 445
HsUsherIb-VIIA LEQEEYDLESIDWL--HIEFTDNQDALDM--IANKPMNIISLID---------- 445
Dm10A-XV LEQAEYARERLEWT--PLAWDDNLPVIHL--LAKKPVGICHLLD---------- 440
Mm-MYO15 EEQEEYIREQMDWR--EIAFADNQPCINL--ISLKPYGILRILD---------- 439
Hs-MYO15 EEQEEYIREQIDWQ--EITFADNQPRINL--ISLKPYGILRILD---------- 439
DmMhcl-PDZ APRDRYAQELVEIEMDLASECHPGPLIS-LIDKAPQNHVVRSSQRDLREHDRRGMLWLLD 476
Hs-PDZ QELERYKEENIELAFDDLEPPTDDSVAAFTVDQASHQSLVRSLAR---TDEARGLLWLLE 440
Mm-PDZ QELERYKEDNIELAFDDLEPVADDSVAA--VDQASH--LVRSLAH---ADEARGLLWLLE 474
Dm29D SSVESCRDEGIVCD-TEVDYVDNVPCIDLISSLR--TGLLSMLD---------- 479
CONSENSUS LEQEEY-REGI-W----I-Y-DNQ--IDLI-------GIL-LLD----------

Actin binding

```
Ac-IB              TMHAQSDGMDGKFLQKCQGGFPSHLHFRGMN-------------NAFSIKHYAGEV--T 484
Ac-IC              ----FPKGTDDKFREKLLGAFPTHAHLAATSQ------------PDEFVIKHYAGDV--V 476
ScMyo3-IA          TAHADSNAADQAFAQRLN-LFNSNPYFELRA--------------NKFVIKHYAGDV--T 494
ScMyo5-IB          TAHADSNAADQAFAQRLN-LFTTNPHFDLRS--------------NKFVIKHYAGDV--T 494
Cehum5-IA          SI---GNVTDKVFLGELDKKLKSHKHYTSRNLKQSD---KSMGF-EEFKITHYAGDV--T 505
CehumI-IB          QNHGQREGVDRQLLTTLSKSFAGHPHFGPGS--------------DSFVIKHYAGDV--T 494
Dm31D-IA           SV---GKVTDDTLLGAMDKNLSKHPHYTSRQLKPTD---KELKHREDFRITHYAGDV--I 500
Dm61F-IB           RP---GEPTDKTFLEKLTQKLAQHHHYVCHEKAPAH--IKKIMLRDEFRLVHYAGEV--T 501
Dm95E              ---EPHLNSNDALLLRVQQCCAGHPNFMTTGSNS-----------MCFQIRHYASVVN-- 475
Rc-IB              RP---GEATDMTFLEKLEDTVKNHPHFVTHKLGDQK--TRKVLGRDEFRLLHYAGEVNFT 523
Gg-Bb              RP---GTVNEDTFITKLNQIFASHKRYESKETLNAKHVTDVSLPLRCFRIHHYAGKVTYF 522
Gg-IB              TMHATGEGADQTLLQKLQAAVGTHEHFNSWS--------------SGFVIHHYAGKV--S 492
Rnmyr1-IA          RP---GTVTDETFLEKLNQVCATHQHFESRMSKCSRFLNDTTLPHSCFRIQHYAGKVLY- 505
Rnmyr2-IB          RP---GEATDLTFLEKLEDTIKHHPHFLTHKLADQK--TRKSLDRGEFRLLHYAGEV--T 505
Rnmyr3-IC          TMHAVGEGADQTLLQKLQMQIGSHEHFNSWN--------------QGFIIHHYAGKV--S 492
Rnmyr4-ID          NV---GKVTDGMFLEALNSKLGKHGHFSSRKTCASD---KILEFDRDFRIRHYAGDV--V 503
ScMYO1-IIA         EEAVLPKSLMESFYSKLISTWDQNSSKFKRSALKN-----------GFILKDYAG--DVN 510
ScMYS1-IIB         EEAVLPKSTDESFYSKLISTWDQNSSKFKRSRLKN-----------GFILKHYAG--DVE 512
Dd-nmII            EQSVFPNATDNTLITKLHSHFSKKNAKYEEPRFSK----------TEFGVTHYAG--QVM 491
Cemyo3-IIA         EECIVPKATDMTYAQKLLDQHLGKHPNFQKPKPPKGKQ-----GDAHFAIVHYAG--TVR 505
Cemyo4-IIB         EECIVPKATDLTLASKLVDQHLGKHPNFEKPKPPKGKQ-----GEAHFAMRHYAG--TVR 505
Cemyo2-IIC         EECIVPKATDLTLAQKLIDQHLGKHPNFEKPKPPKGKQ-----AEAHFAMRHYAG--TVR 511
Cemyo1-IID         EECIVPKATDMTLAQKLTDQHLGKHPNFEKPKPPKGKQ-----GEAHLAMRHYAG--TVR 502
Dm-mhcII           EESMFPKATDQTFSEKLTNTHLGKSAPFQKPKPPKPGQ-----QAAHFAIAHYAG--CVS 500
Dmzip-nmII         EECWFPKATDKTFVDKLVSAH-SMHPKFMKT-DFR--------GVADFAIVHYAG--RVD 540
Xl-nmIIB           EECWFPKATDKTFVDKLVQEQ-GTHSKFQKPRQLK--------DKADFCIIHYAG--RVD 515
Gg-FSkII           EECMFPKATDTSFKNKLYDQHLGKSNNFQKPKPAK-GK-----AEAHFSLVHYAG--TVD 502
Gg-nmII            EECWFPKATDKSFVEKVVQEQ-GTHPKFQKPKQLK--------DKADFCIIHYAG--KVD 496
Gg-smII            EECWFPKATDTSFVEKLIQEQ-GNHAKFQKSKQLK--------DKTEFCILHYAG--KVT 506
Hs-FSkEII          EECMFPKATDTSFKNKLYDQHLGKSNNFQKPKVVK-GR-----AEAHFSLIHYAG--TVD 501
Hs-nmIIA           EECWFPKATDKSFVEKVMQEQ-GTHPKFQKPKQLK--------DKADFCIIHYAG--KVD 496
Hs-CaA             EECMFPKATDMTFKAKLYDNHLGKSNNFQKPRNIK-GK-----PEAHFSLIHYAG--TVD 502
DmNinaC-III        DASRSCQD-----QNLIMDRVSEKHSQFVKKHTA----------TEISVAHYTGRII-- 490
Lp-III             KATQYQRKGESGWMESLQEADSEELVEFSNVNGK----------PIVSVKHIFRKVS-- 456
HsMYO3A            EESRFPKATDQTLVEKFEGN-LKSQYFWRPKRME----------LSFGIHHYAGKVL-- 498
ScMyo2p-Va         EESRLPAGSDESWTQKLYQTLDKSPTNKVFSKPRFGQT---------KFIVSHYA-LDVA 500
ScMyo4p-Vb         EESRLPSGSDESWASKLYSAFNKPPSNEVFSKPRFGQT---------KFIVSHYA-VDVE 500
Cehum2-V           EQCKRLNGSDADWLSQLQNST-ELKRNPQLAFPKVRSN---------DFIVRHFA-ADVT 507
SpMyo-Va           EECMLPKGSDENWCNKLYN---KLTAHGHFSKPRTSRT---------AFLVHHFA-DKVE 487
Dmdidum-Va         EECRMPKGSDESWAGKLIG---KCNKFPHFEKPRFGTT---------SFFIKHFS-DTVE 496
Ggp190-Va          EECKMPKGSDDTWAQKLYNT--HLNKCALFEKPRLSNK---------AFIIKHFA-DKVE 488
Mmdilute-Va        EECKMPKGTDDTWAQKLYNT--HLNKCALFEKPRMSNK---------AFIIKHFA-DKVE 488
Rnmyo5-Va          EECKMPKGTDDTWAQKLYNT--HLNKCALFEKPRMSNK---------AFIIKHFA-DKVE 488
Rnmyr6-Vb          EECKVPKGTDQNWAQKLYER--HSN-SQHFQKPRMSNT---------AFIVIHFA-DKVE 488
HsMYO5A            EECKMPKGTDDTWAQKLYNT--HLNKCALFEKPRLSNK---------AFIIQHFA-DKVE 488
HsMYO5C            EECLLPHGTDENWPQKLYNN--FVNRNPLFEKPRMSNT---------SFVIQHFA-DKVE 488
SpMyoVI            EESKLPKASANNFTHNVHQKHKTHFRLAIPRKSKLTAHRNLRDDEG--ILIKHFA-GAVC 525
Cehum3-VI          EEAKLPRATFQHFTQRAHESNKNHFRLDAPRKSKVKSHREMRDDEG--LLIRHYA-GTVC 528
Dmjar-VI           EESKLPKPSYSHFTAEVHKSWANHYRLGLPRSSRLKAHRTLRDEEG--FLVRHFA-GAVC 529
Ms-VIA             EENRLPQPTDQHFTDTVHNKHKNHFRLTIPRKSKLAVHRNVRDDEGFTFIIRHFA-GAVC 547
Ms-VIB             EENRLPQPSDQHFALAVHNKHKDHFRLTVPRKSKLTIHRNLRDDEGFTFIIRHFA-GAVC 547
GgMyoVI            EENRLPQPSDQHFTSVVHQKHKDHFRLSIPRKSKLAVHRNVRDDEG--FIIRHFA-GAVC 529
MmWaltzerVI        EENRLPQPSDQHFTSVVHQKHKDHFRLTIPRKSKLAVHRNLRDDEG--FIIRQLCRGRVL 531
HsMYO6             EENRLPQPSDQHFTSAVHQKHKDHFRLTIPRKSKLAVHRNIRDDEG--FIIRHFA-GAVC 529
Cehum6-VII         EESIFPKGTDKTMLLKLHST---HGRNELYLQPKSELQR--------AFGVTHFA-GNVF 489
Dmck-VIIA          EEARFPKGTDQTMLAKLHKT---HGSHKNYLKPKSDINT--------SFGLNHFA-GVVF 489
Dm28B-VIIB         EESKFPKGTDQTLLEKLHVQ---HGNRSIYVKGKTTQTS--------LFGIRHYA-GVVM 491
DrMyo-VIIA         EESKFPKGTDTTMLNKLNSQ---HKLNTNYIPPKNTYET--------QFGIQHFA-GVVY 493
Mmshaker-VIIA      EESKFPKGTDATMLHKLNSQ---HKLNANYVPPKNSHET--------QFGINHFA-GVVY 493
HsUsherIb-VIIA     EESKFPKGTDTTMLHKLNSQ---HKLNANYIPPKNNHET--------QFGINHFA-GIVY 493
Dm10A-XV           DESNFPRATDLSFLEKCHYN---HALSELYARPRIGAQ---------EFGVTHYA-GQVW 487
Mm-MYO15           DQCCFPQATDHTFLQKCHYH---HGANPLYSKPKMPLP---------EFTIKHYA-GKVT 486
Hs-MYO15           DQCCFPQATDHTFLQKCHYH---HGANPLYSKPKMPLP---------EFTIKHYA-GKVT 486
DmMhcl-PDZ         EEAIYPNSNDDTFLERLFSHYGDREHH------SLLRKCA---GPRQFVLHHLQGTNPVL 527
Hs-PDZ             EEALVPGASEDTLLERLFSYYGPQEGDKKGQSFTPLLHSS---KPHHFLLGHSHGTNWVE 497
Mm-PDZ             EEALVPGATEDALLDRLFSYYGPQEGDKKGQ--SPLLRSS---KPRHFLLGHSHGTNWVE 529
Dm29D              -AECSVRGTAESYVTKLKVQHRSSTRLETKPTAEPHDP-------RMFLIRHFAG--RVE 529
CONSENSUS          EE---PKATD-TFL-KL--------------K--------------F-IIHYAG---V-
```

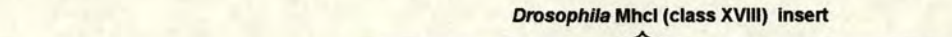

Loop 2
Ac-IB YEAEGFCE-KNKDTLFDDLIAVIQESE--NRLLVSWFP-EDTKQLQ---------- 526
Ac-IC YNVDGFCD-KNKDLLFKDLIGLAECTS--STFFAGLFP-EAKEVATS---------- 519
ScMyo3-IA YDINGITD-KNKDQLQKDLIELIGTTT--NTFLSTIFP-DDVDKDS---------- 536
ScMyo5-IB YDIDGITD-KNKDQLQKDLVELIGTTT--NTFLATIFP-DTVDRES---------- 536
Cehum5-IA YSVMGFMD-KNKDTLFQDLKRLLYHSK--NRLVKSLFPDGSKSMAEV---------- 549
CehumI-IB YNVDGFCD-RNRDVLYPDLILLMQKSS--RPFIQALFP-ENVAASA---------- 536
Dm31D-IA YNINGFIE-KNKDTLYQDFKRLLHNSK--DANLSEMWPEGAQDIKKT---------- 544
Dm61F-IB YSVNGFLD-KNNDLLFRDLKETLSKAG--NGIVRSCFPEKELR---S---------- 542
Dm95E YSIHRFLE-KNSDMLPKYISAAFYQSK--LSLVQSLFPEGNPRR-QV---------- 518
Rc-IB YSVAGFLD-KNNDLLFRNLKEVMCDSG--NPIAHQCFNRSELT---D---------- 564
Gg-Bb TNVTGFIE-KNNDLLFRDLSQAMWAAR--HTLLRSLFPEGDPQR-PS---------- 565
Gg-IB YDVNGFCE-RNRDVLFTDLIELMQSSE--YAFIRMLFP-EKLDADK---------- 534
Rnmyr1-IA -QVEGFVD-KNNDLLYRDLSQAMWKAG--HALIKSLFPEGNPAK-VN---------- 547
Rnmyr2-IB YSVTGFLD-KNNDLLFRNLKETMCSST--NPIMAQCFDKSELS---D---------- 546
Rnmyr3-IC YDMDGFCE-RNRDVLFMDLIELMQSSE--LPFIKSLFP-ENLQADK---------- 534
Rnmyr4-ID YSVIGFID-KNKDTLFQDFKRLMYNSS--NPVLKNMWPEGKLSITEV---------- 547
ScMYO1-IIA TLWKAGYP-KT-DPLNDNLLSLLSSSQ--NDIISKLFQPEGGKNLLVCGVEANISNQEV- 565
ScMYS1-IIB YTVEGWLS-KNKDPLNDNLLSLLSSSQ--NDIISKLFQPEGEK-SSSAGVEANISNQEV- 567
Dd-nmII YEIQDWLE-KNKDPLQQDLELCFKDSS--DNVVTKLFNDPNIASRAKKGAN--------- 539
Cemyo3-IIA YNATNFLE-KNKDPLNDTAVALLKHSTD-NSLMLDIWQDYQTQEEAAEAAKAGQTAGG-- 561
Cemyo4-IIB YNCLNWLE-KNKDPLNDTVVSAMKQSKG-NDLLVEIWQDYTTQEEAAAKAKEGGGGGK-- 561
Cemyo2-IIC YNCLNWLE-KNKDPLNDTVVTVMKASKE-HALIVEVWQDYTTQEEAAAAAAKGTAGAK-- 567
Cemyo1-IID YNVLNWLE-KNKDPLNDTVVSVMKASKK-NDLLVEIWQDYTTQEEAAAAAKAGGG--R-- 556
Dm-mhcII YNITGWLE-KNKDPLNDTVVDQFKKSQ--NKLLIEIFADHAGQSGGGEQAKGGRG----- 552
Dmzip-nmII YSAAKWLM-KNMDPLNENIVSLLQGSQ--DPFVVNIWKDAE-IVGMAQQA-LTDTQFG-- 593
Xl-nmIIB YKADEWLL-KNMDPLNDNVATLLHQSS--DKFVSELWKDVDRIVGLDQVAGMAETAFGA- 571
Gg-FSkII YNISGWLE-KNKDPLNETVIGLYQKSS--VKTLALLFATYGGEAEG-GGGKKGGK----- 553
Gg-nmII YKADEWLM-KNMDPLNDNIATLLHQSS--DKFVSELWKDVDRIVGLDQVAGMSETALPG- 552
Gg-smII YNASAWLT-KNMDPLNDNVTSLLNQSS--DKFVADLWKDVDRIVGLDQMAKMTESSLPS- 562
Hs-FSkEII YSVSGWLE-KNKDPLNETVVGLYQKSS--NRLLAHLYATFATADAD-SGKKKVAK----- 552
Hs-nmIIA YKADEWLM-KNMDPLNDNIATLLHQSS--DKFVSELWKDVDRIIGLDQVAGMSETALPG- 552
Hs-CaA YNILGWLE-KNKDPLNETVVGLYQKSS--LKLMATLFSSYATADTGDSGKSKGGK----- 554
DmNinaC-III YDTRAFTD-INRDFVPPEMIETFRSSL--DESIMLMFTNQLTKAGNLTMPFEAVQHKDES 547
Lp-III YDATDLVK-KNVEDKTRALTSTMQRSC--DPRIRAIFSSENPSP----------FLSSP 502
HsMYO3A YNASGFLA-KNKTLVPTDIVLLLRSSD--NSVIRQLVNHPLTKTGN--LPHSKLKMLNYQ 553
ScMyo2p-Va YDVEGFIE-KNRDTVSDGHLEVLKAST--NETLINILEGLEKAAKKLEEAKKL------- 550
ScMyo4p-Vb YEVEGFIE-KNRDSVSLGHLDVFKATT--NPIFKQILD-----NRELRSDDAP------- 545
Cehum2-V YSTDGFVE-KNRDAIGEQLLDVVVASK--FPFIRTVIGST---------- 544
SpMyo-Va YESEGFVQ-KNRDQVNDEHLNILMASQ--YEFVAELFRPK-----PEAPKHNHK------ 533
Dmdidum-Va YDVNGFLE-KNRDTVSKELTQVLSESN--MSLAKQVMTLEEIDTLCVDSAKSSTLG---- 549
Ggp190-Va YQCEGFLE-KNKDTVYEEQIKVLKSSKK-FKLLPELFQDEEKAISPTSATPSGRVPL--- 543
Mmdilute-Va YQCEGFLE-KNKDTVFEEQIKVLKSSK--FKMLPELFQDDEKAISPTSATSSGRTPL--- 542
Rnmyo5-Va YQCEGFLE-KNKDTVFEEQIKVLKSSK--FKMLPELFQDDEKAISPTSATSSGRTPL--- 542
Rnmyr6-Vb YLSDGFLE-KNRDTVYEEQINILKASK--FPLVADLFRDDEDSVPATNTAKSRSS----- 540
HsMYO5A YQCEGFLE-KNKDTVFEEQIKVLKSSK--FKMLPELFQDDEKAISPTSATSSGRTPL--- 542
HsMYO5C YKCEGFLE-KNRDTVYDMLVEILRASK--FHLCANFFQEN-----PTPPSPFG------- 533
SpMyoVI YQTSEFIE-KNNDALHASLEVVIRDCK--DPFITSLFPKNDKEPSKGRGGQNVQ------ 576
Cehum3-VI YETRYFVE-KNNDQLHNSLEMLIEQSS--FPLVVSLFTSEATGAVKTGG---------- 574
Dmjar-VI YNTEQFIE-KNNDALHASLEGLVQECD--NPLLQTLFPSGSSTSVRG---------- 573
Ms-VIA YETTNFVE-KNNDALHMSLECLVSESK--DRFVRELFE-NSNNSKDVKQKPEFT------ 597
Ms-VIB YETCRFVE-KNNDALHMSLESLVCESK--DKFVRELFE-NSNTTKDSKQKAGFT------ 597
GgMyoVI YETTQFVE-KNNDALHMSLESLICESK--DKFVRQLFESNTNNNKDPKQKAG-------- 578
MmWaltzerVI RRQPQYGGGKNNDALHMSLESLICESR--DKFIRALFESSTNNSKDTKQKAG-------- 581
HsMYO6 YETTQFVE-KNNDALHMSLESLICESR--DKFIRELFESSTNNNKDTKQKAG-------- 578
Cehum6-VII YNTRGFLE-KNRDSFSADLSVLISSSK--MPFLARLFD-DIEYDTSSRKK---------- 535
Dmck-VIIA YDTRGFLD-KNRDTFSPDLLHLVSQST--NKFLRQIFAQDIEMGAETRKRT--------- 537
Dm28B-VIIB YNPLGFLE-KNRDSFSGDLRTLVQRST--NKYLVDIFPHEMPM--DTAKKQ--------- 537
DrMyo-VIIA YETRGFLE-KNRDTLHGDIIQLVHSSK--NKFIKQIFQADVAMGAETRKRS--------- 541
Mmshaker-VIIA YESQGFLE-KNRDTLHGDIIQLVHSSR--NKFVKQIFQADVAMGAETRKRS--------- 541
HsUsherIb-VIIA YETQGFLE-KNRDTLHGDIIQLVHSSR--NKFIKQIFQADVAMGAETRKRS--------- 541
Dm10A-XV YCVDGFLD-KNRDALRGDVLELLASSR--LPLVGELTKQLRAQRDAGKTLPKGSNG---- 540
Mm-MYO15 YQVHKFLD-KNHDQVRQDVLDLFVHSR--TRVVAHLFSSHAAQTAP-PRLGKSSS----- 537
Hs-MYO15 YQVHKFLD-KNHDQVRQDVLDLFVRSR--TRVVAHLFSSHAPQAAP-QRLGKSSS----- 537
DmMhcl-PDZ YAVDGWVRHSREHPGIRNAVSLLQDSSR-EEINRLYIGSLTRGSGAMVFCGSFAGLEG-- 584
Hs-PDZ YNVTGWLNYTKQNPATQNVPRLLQDSQK-KIISNLFLFTGRAGS-ATVLSGSIAGLEGGS 555
Mm-PDZ YNVAGWLNYTKQNPATQNAPRLLQDSQK-KIISNLFL--GRAGS-ATVLSGSIAGLEGGS 585
Dm29D YDTTDFLD-TNRDVVPDDLVGVFYKHTCNFGFATHLFGSELKALYAQQQAPRGLSFRIS- 587
CONSENSUS Y---GFLE-KNKD-L---LI-LL--S-----L---LF----------


Class V insert
Actin contact helix
Ac-IB ----------KKRPTTAGFKLKT-------SCDALMEALSRCSPHYIRCIK 560
Ac-IC ----------KKKPTTAGFKIKE-------SINILVATLSKCTPHYIRCIK 553
ScMyo3-IA ----------KRRPPTAGDKIIK-------SANELVETLSKAEPSYIRTIK 570
ScMyo5-IB ----------KRRPPTAGDKIIK-------SANDLVETLSKAQPSYIRTIK 570
Cehum5-IA ----------NRRPPTAGFLFKN-------SMSELVKQLAQKEPHYIRCIK 583
CehumI-IB ----------GKRPTTFSTKIRT-------QANTLVESLMKCSPHYVRCIK 570
Dm31D-IA ----------TKRPLTAGTLFQR-------SMADLVVTLLKKEPFYVRCIK 578
Dm61F-IB ----------LKRPETAITQFR--A-----SLNNLMDILMCKEPSYIRCIK 576
Dm95E ----------TKKPSTLSSNIRT-------QLQTLLAIVKHRRSHYVFCIK 552
Rc-IB ----------KKRPETAATQFKNSL-----SKLFTMEILMSKQPSYVRCIK 600
Gg-Bb ----------LKLPPTTGSQFKASV-----ATFTLMKNLYSKNPNYIRCIK 601
Gg-IB ----------KGRPTTAGSKIKK-------QANDLVNTLMKCTPHYIRCIK 568
Rnmyr1-IA ----------LKRPPTAGSQFKASV-----A--TLMKNLQTKNPNYIRCIK 581
Rnmyr2-IB ----------KKRPETVATQFK--M-----SLLQLVEILRSKEPAYIRCIK 580
Rnmyr3-IC ----------KGRPTTAGSKIKK-------QANDLVSTLMKCTPHYIRCIK 568
Rnmyr4-ID ----------TKRPLTAATLFKN-------SMIALVDNLASKEPYYVRCIK 581
ScMYO1-IIA ----------KKSARTSTFKTTSSRHREQQI-------TLLNQLASTHPHFVRCII 604
ScMYS1-IIB ----------KKSARTSTFKTTSSRHREQQI-------TLLNQLASTHPHFVRCII 606
Dd-nmII ----------FITVAAQYKEQLA-------SLMATLETTNPHFVRCII 570
Cemyo3-IIA ----------KRGKSSSFATVSMIYRESLN-------NLMNMLYQTHPHFIRCII 599
Cemyo4-IIB ----------KKGKSGSFMTVSMLYRESLN-------NLMTMLNKTHPHFIRCII 599
Cemyo2-IIC ----------KKGKSGSFMTVSMLYRESLN-------KLMTMLHSTHPHFIRCII 605
Cemyo1-IID ----------KGGKSGSFMTVSMMYRESLN-------KLMTMLHKTHPHFIRCII 594
Dm-mhcII ----------KKG--GGFATVSSAYKEQLN-------SLMTTLRSTQPHFVRCII 588
Dmzip-nmII ----------ARTRKGMFRTVSHLYKEQLA-------KLMDTLRNTNPNFVRCII 631
Xl-nmIIB ----------AYKTKKGMFRTVGQLYKESLA-------KLMATLRNTNPNFVRCII 610
Gg-FSkII ----------KKG--SSFQTVSALFRENLN-------KLMANLRSTHPHFVRCII 589
Gg-nmII ----------AFKTRKGMFRTVGQLYKEQLA-------KLMATLRNTNPNFVRCII 591
Gg-smII ----------ASKTKKGMFRTVGQLYKEQLT-------KLMTTLRNTNPNFVRCII 601
Hs-FSkEII ----------KKG--SSFQTVSALFRENLN-------KLMSNLRTTHPHFVRCII 588
Hs-nmIIA ----------AFKTRKGMFRTVGQLYKEQLA-------KLMATLRNTNPNFVRCII 591
Hs-CaA ----------KKG--SSFQTVSALHRENLN-------KLMTNLRTTHPHFVRCII 590
DmNinaC-III ERKSYALNTLSAGCISQVNNLRTLAANFRFTCLT------LLKMLSQNANLGVHFVRCIR 601
Lp-III RRSSIQENMLLP--------ERTVTDSLHSALSS------VLNLASTEDPP--HLILCMR 546
HsMYO3A MRTSEKLINLAKGDTGEATRHARETTNMKTQTVASYFRYSLMDLLSKMVVGQPHFVRCIK 613
ScMyo2p-Va ----ELEQAGSKKPGPIRTVNRKPTLGSMFKQSLI------ELMNTIN-STNVHYIRCIK 599
ScMyo4p-Vb ----EEQNTEKKIMIPARLSQKKPTLGSMFKKSLG------ELMAIIN-STNVHYIRCIK 594
Cehum2-V ---APTSVSSSSSSSTPGKRTIKKTVASQFRDSLK------ELMSVLC-STRPHYVRCIK 594
SpMyo-Va ----RGSVKPMVAPVSR-TKTFKRSVGSQFRDSLT------YLMMKLN-STTPHYVRCIK 581
Dmdidum-Va ----GRVVISAGRKQVVPSKQHRKTVGSQFQESLA------SLISTLH-ATTPHYVRCIK 598
Ggp190-Va ---SRTPVKPAKARPGQTSKEHKKTVGHQFRNSLH------LLMETLN-ATTPHYVRCIK 593
Mmdilute-Va ---TRVPVKPTKGRPGQTAKEHKKTVGHQFRNSLH------LLMETLN-ATTPHYVRCIK 592
Rnmyo5-Va ---TRVPVKPTKGRPGQTAKEHKKTVGLQFRNSLH------LLMETLN-ATTPHYVRCIK 592
Rnmyr6-Vb ---SKINVRSSRPLMKAPNKEHKKSVGYQFRTSLN------LLMETLN-ATTPHYVRCIK 590
HsMYO5A ---TRTPAKPTKGRPGQMAKEHKKTVGHQFRNSLH------LLMETLN-ATTPHYVRCIK 592
HsMYO5C ---SMITVKSAKQVIKPNSKHFRTTVGSKFRSSLY------LLMETLN-ATTPHYVRCIK 583
SpMyoVI ----------KLAFDSVGNKFKVQLN------QLMEKLR-STGSSFVRCIK 610
Cehum3-VI ----------RLKAVSVGAKFKSQLS------SLLDKLN-NTGTHFVRCVK 608
Dmjar-VI ----------KLNFISVGSKFKTQLG------ELMEKLE-QNGTNFIRCIK 607
Ms-VIA ----------KLSFISVGNKFKTQLN------ILLEQARCSTGSSFIRCIK 632
Ms-VIB ----------KLGFISVGNKFKTQLN------LLLEKLR-STGSSFIRCVK 631
GgMyoVI ----------KLSFISVGNKFKTQLN------LLLEKLH-STGSSFIRCIF 612
MmWaltzerVI ----------KLSFISVGNKFKTQLN------LLLDKLR-STGASFIRCIK 615
HsMYO6 ----------KLSFISVGNKFKTQLN------LLLDKLR-STGASFIRCIK 612
Cehum6-VII ----------VTVGNQFRRSLE------QLMSQLT-QTHPFFIRCIK 565
Dmck-VIIA ----------PTLSTQFRKSLD------ALMKTLS-SCQPFFIRCIK 567
Dm28B-VIIB ----------PTLCVKFRNSLD------MLMRTLS-QAHPYFIRCIK 567
DrMyo-VIIA ----------PTLSSQFKRSLE------LLMRTLS-VCQPFFVRCIK 571
Mmshaker-VIIA ----------PTLSSQFKRSLE------LLMRTLG-ACQPFFVRCIK 571
HsUsherIb-VIIA ----------PTLSSQFKRSLE------LLMRTLG-ACQPFFVRCIK 571
Dm10A-XV ----------RFVTMKPRTPTVAARFADSLQ------QLLQSMG-RCHPWFVRCIK 579
Mm-MYO15 ----------ITRLYKAHTVAAKFQQSLL------DLVEKME-RCNPLFVRCLK 574
Hs-MYO15 ----------VTRLYKAHTVAAKFQQSLL------DLVEKME-RCNPLFMRCLK 574
DmMhcl-PDZ TQSLRRVSSIRRSFTT--AGVKRNSIMLQVKFTVD-----GIID--TLRRTGTHFVHCYL 635
Hs-PDZ QLALRRATSMRKTFTTGMVAVKKKSLCIQMKLQVD-----ALIDFTTIKKSKLHFVHCFL 610
Mm-PDZ QLALRRATSMRKTFTTGMAAVKKKSLCIQIKLQVD-----ALID--TIKRSKMHFVHCFL 638
Dm29D ---------PTSHSDLLNGDEPVSTLTQDFHTRLDN-------LLRTLVHARPHFVRCIR 631
CONSENSUS ----------TVG--FK--L---------LM-LL--T-PHFVRCIK


Class XVIII insert
Ac-IB PNDNKAYHDWDATRTKHQVQYLGLL----------ENVR 589
Ac-IC PNEKKAANAFNNSLVLHQVKYLGLL----------ENVR 582
ScMyo3-IA PNQTKSPNDYDDHQVLHQVKYLGLQ----------ENVR 599
ScMyo5-IB PNETKSPNDYDDRQVLHQIKYLGLQ----------ENVR 599
Cehum5-IA PNEEKNSNVFDLERVEHQVRYLGLL----------ENVR 612
CehumI-IB PNETKRPNDWEESRVKHQVEYLGLR----------ENIR 599
Dm31D-IA PNDLKSSTVFDEERVEHQVRYLGLL----------ENLR 607
Dm61F-IB PNDLQTANVFNDELVLHQVKYLGLM----------ENLR 605
Dm95E PNEGKQPHQFDMALVQHQVRYMSLM----------PLVH 581
Rc-IB PNDAKQPARFDEVLIRHQVKYLGLI----------ENVR 629
Gg-Bb PNDTKTAMLFTPDLVLAQVRYLGLM----------ENVR 630
Gg-IB PNETKKPRDWEESRVKHQVEYLGLK----------ENIR 597
Rnmyr1-IA PNDKKAAHIFSESLVCHQIRYLGLL----------ENVR 610
Rnmyr2-IB PNDAKQPGRFDEVLIRHQVKYLGLM----------ENLR 609
Rnmyr3-IC PNETKKPKDWEESRVKHQVEYLGLK----------ENIR 597
Rnmyr4-ID PNDKKSPQIFDDERCRHQVEYLGLL----------ENVR 610
ScMYO1-IIA PNNVKKVKTFNRSLILDQLRCNGVL----------EGIR 633
ScMYS1-IIB PNNVKKVKTFNRRLILDQLRCNGVL----------EGIR 635
Dd-nmII PNNKQLPAKLEDKVVLDQLRCNGVL----------EGIR 599
Cemyo3-IIA PNEKKASGVIDSALVLNQLTCNGVL----------EGIR 628
Cemyo4-IIB PNEKKQSGMIDAALVLNQLTCNGVL----------EGIR 628
Cemyo2-IIC PNEKKASGVIDAGLVLNQLTCNGVL----------EGIR 634
Cemyo1-IID PIEKKQSGMIDAALVLNQLTCNGVL----------EGIR 623
Dm-mhcII PNEMKQPGVVDAHLVMHQLTCNGVL----------EGIR 617
Dmzip-nmII PNHEKRAGKIDAPLVLDQLRCNGVL----------EGIR 660
Xl-nmIIB PNHEKRAGKLDPHLVLDQLRCNGVL----------EGIR 639
Gg-FSkII PNETKTPGAMEHELVLHQLRCNGVL----------EGIR 618
Gg-nmII PNHEKKAGKLDPHLVLDQLRCNGVL----------EGIR 620
Gg-smII PNHEKRAGKLDAHLVLEQLRCNGVL----------EGIR 630
Hs-FSkEII PNETKTPGAMEHSLVLHQLRCNGVL----------EGIR 617
Hs-nmIIA PNHEKKAGKLDPHLVLDQLRCNGVL----------EGIR 620
Hs-CaA PNERKAPGVMDNPLVMHQLRCNGVL----------EGIR 619
DmNinaC-III ADLEYKPRSFHSDVVQQQMKALGVL----------DTVI 630
Lp-III PQKKELINDYDSKSVQIQLHALNVL----------ETIL 575
HsMYO3A PNSERQARKYDKEKVLLQLRYTGIL----------ETAR 642
ScMyo2p-Va PNADKEAWQFDNLMVLSQLRACGVL----------ETIR 628
ScMyo4p-Vb PNSEKKPWEFDNLMVLSQLRACGVL----------ETIR 623
Cehum2-V PNDSKISFDFEPKRAIQQLRACGVL----------ETVR 623
SpMyo-Va PNDYKLPFTFEPKRAVEQLRACGVL----------ETIR 610
Dmdidum-Va PNDDKVAFKWETAKIIQQLRACGVL----------ETVR 627
Ggp190-Va PNDFKFPFTFDEKRAVQQLRACGVL----------ETIR 622
Mmdilute-Va PNDFKFPFTFDEKRAVQQLRACGVL----------ETIR 621
Rnmyo5-Va PNDFKFPFTFDEKRAVQQLRACGVL----------ETIR 621
Rnmyr6-Vb PNDEKLPFHFDPKRAVQQLRACGVL----------ETIR 619
HsMYO5A PNDFKFPFTFDEKRAVQQLRACGVL----------ETIR 621
HsMYO5C PNDEKLPFEFDSKRIVQQLRACGVL----------ETIR 612
SpMyoVI PNGKMVDSLFEGGQILSQLQCSGMV----------TVLD 639
Cehum3-VI PNSQMKAWHFDGSAILGQLQCAGMA----------SVLR 637
Dmjar-VI PNSKMIDRQFEGSLALAQLKCSGTI----------SVLE 636
Ms-VIA PNLKMVSHQFEGAQILSQLQCSGMFT----------VSVLD 663
Ms-VIB PNLKMVSHQFEGALILSQLQCSGMVF----------TSVLD 662
GgMyoVI PNLKMTSHHFEGGQILSQLQCSGMV----------SVLD 641
MmWaltzerVI PNLKMASHHFEGAQILSQLQCSGMV----------SVLD 644
HsMYO6 PNLKMTSHHFEGAQILSQLQCSGMV----------SVLD 641
Cehum6-VII PNEMKRALVMDRDLVLRQLRYSGMM----------ETIK 594
Dmck-VIIA PNELKKPMMFDRGLCCRQLRYSGMM----------ETIR 596
Dm28B-VIIB PNEYKEPKNFDKELCVRQLRYSGMM----------ETAR 596
DrMyo-VIIA PNEYKKPMLFDRELCVRQLRYSGMM----------ETIR 600
Mmshaker-VIIA PNEFKKPMLFDRHLCVRQLRYSGMM----------ETIR 600
HsUsherIb-VIIA PNEFKKPMLFDRHLCVRQLRYSGMM----------ETIR 600
Dm10A-XV PNQEKHALRMDMPCVLQQLRYLGML----------DTIQ 608
Mm-MYO15 PNHKKEPGLFEPDVMMAQLRYSGVL----------ETVR 603
Hs-MYO15 PNHKKEPGLFEPDVVMAQLRYSGVL----------ETVR 603
DmMhcl-PDZ LQHN--AGKHTKYTANGSPSSAAGQVSSEE------EMVNVPLLR--SQLRGSQVLEAAR 685
Hs-PDZ PVAEGWAGEPRSASSRRVSSSSELDLPSGDHCEAGLLQLDVPLLRFTTQLRGSRLLDAMR 670
Mm-PDZ PVAEGWPGEPRSASSRRVSSSSELDLPPGDPCEAGLLQLDVSLLR--AQLRGSRLLDAMR 696
Dm29D SNGTEAARSFDRATVVRQIRSLQVL----------ETVN 660
CONSENSUS PNE-K----FD--LVL-QLR--GVL----------E-IR

Loop 3

```
Ac-IB                VRRAGFAYRAEFDR----FLRRYKKLSP-KTWGIWGEWS-GAPKDGCQTLLNDLGLDT-S 642
Ac-IC                IRRAGYAYRQSYDK----FFYRYRVVCP-KTW---SGWN-GDMVSGAEAILNHVGMSLGK 633
ScMyo3-IA            IRRAGFAYRQTFEK----FVERFYLLSPDCSYAGDYTWD-GDTLEAVKLILRDAMIPE-K 653
ScMyo5-IB            IRRAGFAYRQVFEK----FVERFYLLSPHCSYAGDYTWQ-GDTLDAVKYILQDSSIPQ-Q 653
Cehum5-IA            VRRAGFAHRMPYDRF----VNRYKLICA-STWP-NPRRG-QQLKDSCMQILESAGLA--Q 663
CehumI-IB            VRRAGFAYRRAFDK----FAQRYAIVSP-QTWP---CFQ-GDQQRACEIICDSVHMEK-N 649
Dm31D-IA             VRRAGFVHRQRYDKF----LLRYKMISQ-YTWP-NFRAG-SD-RDGVRVLIEEKKFA--Q 657
Dm61F-IB             VRRAGFAYRRTYE--LF--LERYKSLSK-STWP-NYKGP-GGPKAGVQQLVKDLGWDE-E 657
Dm95E                LCRTGHCYHLLHVK----FFHRYKLLNS-LTWP--HFHG-GSQVEGIALIIRNLPLPS-A 632
Rc-IB                VRRAGFAYRRKYE--IFTFLQRYKSLCP-DTWP-NWDG---RAMDGVAVLVKSLGYKP-E 681
Gg-Bb                VRRAGYAFRQLYQ--FTPFLERYKMLSR-KTWP-RWTG---GDREGAEVLLAELKFPP-E 682
Gg-IB                VRRAGFAYRRIFHK----FLQRYAILTP-ETWP---SWR-GDERQGVQHLLRSVNMDP-D 647
Rnmyr1-IA            VRRAGYAFRQAY----EPCLERYKMLCK-QTWP-HWKG---PARSGVEVLFNELEIPV-E 660
Rnmyr2-IB            VRRAGFAYRRKYE--AF--LQRYKSLCP-ETWP-VWTG---RPQDGVAVLVRHLGYKP-E 659
Rnmyr3-IC            VRRAGYAYRRVFQK----FLQRYAILTK-ATWP---VWR-GDEKQGVLHLLQSVNMDS-D 647
Rnmyr4-ID            VRRAGFAFRQTYEKF----LHRYKMISE-FTWP-NHDLP-SD-KEAVKKLIERCGFQ--D 660
ScMYO1-IIA           LAREGYPNRIAFQEFFQRYRILYPENSTTTTFSSKLKASTKQNCRFLLTSLQ----LDTK 689
ScMYS1-IIB           LAREGYPNRIAFQEFFQRYRILYPENSTTTTFSSKLKASTKQNCEFLLTSLQ----LDTK 691
Dd-nmII              ITRKGFPNRIIYADFVKRYYLLAPNVPRDA--EDSQKA-----TDAVLKHLN----IDPE 648
Cemyo3-IIA           ICRKGFPNRMLYPDFKHRYAILAADAAKE---SDPKKA-----SVGILDKISVDGNLTDE 680
Cemyo4-IIB           ICRKGFPNRTLHPDFVQRYAILAAKEAKS--DDDKKKC-----AEAIMSKLVNDGSLSEE 681
Cemyo2-IIC           ICRKGFPNRTLHPDFVQRYALLAADESIIG-KTDAKKG-----SALMLARLVKEKKLEEE 688
Cemyo1-IID           ICRKGFPNRTQHPDFVQRYAILAAKEAKS--SDDMKTC-----AGAILQALINQKQLNDE 676
Dm-mhcII             ICRKGFPNRMMYPDFKMRYQILNPAGIVG--VDDPKKC-----GSIILESTA----LDPD 666
Dmzip-nmII           ICRQGFPNRIPFQEFRQRYELLTPNVIPKG-FMDGKKA-----CEKMIQALE----LDSN 710
Xl-nmIIB             ICRQGFPNRIVFQEFRQRYEILTPNSIPRG-FMDGKQA-----CERMIRSLE----LDPN 689
Gg-FSkII             ICRKGFPSRVLYADFKQRYRVLNASAIPEGQFMDSKKA-----SEKLLGSID----VDHT 669
Gg-nmII              ICRQGFPNRVVFQEFRQRYEILTPNAIPKG-FMDGKQA-----CVLMIKALE----LDSN 670
Gg-smII              ICRQGFPNRIVFQEFRQRYEILAANAIPKG-FMDGKQA-----CILMIKALE----LDPN 680
Hs-FSkEII            ICRKGFPNRILYGDFKQRYRVLNASAILEGQFIDSKKA-----CEKLLASID----IDHT 668
Hs-nmIIA             ICRQGFPNRVVFQEFRQRYEILTPNSIPKG-FMDGKQA-----CVLMIKALE----LDSN 670
Hs-CaA               ICRKGFPNRILYGDFRQRYRILNPVAIPEGQFIDSRKG-----AEKLLSSLD----IDHN 670
DmNinaC-III          ARQKGFSSRLPFDE----FLRRYQFLAFDFDEP-----V-EMTKDNCRLLFLRLKMEG-- 678
Lp-III               IRQFGFARRISFVD----FLNRYQYLAFDFNEN-----V-ELTKENCRLLLLRLKMDG-- 623
HsMYO3A              IRRLGFSHRILFAN----FIKRYYLLCYKSSEE-----P-RMSPDTCATILEKAGLDN-- 690
ScMyo2p-Va           ISCAGFPSRWTFEEFVLRYYILIPHEQWDLIFKKKETTE-EDIISVVKMILDATVKDK-S 686
ScMyo4p-Vb           ISCAGFPSRWTFDEFVQRYFLLTDYSLWSGILYNPDLPK-EAIVNFCQSILDATISDS-A 681
Cehum2-V             ISAAGFPSRYPYEEFARRYRVIYTKEAALWR----------DKPKQFAELACQQCLEE-G 672
SpMyo-Va             ISAAGYPSRWSYSEFFTRYRVLMVRKEINKK----------DVRGTSEKTLKRLIVDP-D 659
Dmdidum-Va           ISAAGFPSRWLYPDFYMRYQLLVYRSKLDKN----------DMKLSCRNIVMKWIQDE-D 676
Ggp190-Va            ISAAGFPSRWTYQEFFSRYRVLMKQKDVLS-----------DRKQTCKNVLEKLILDK-D 670
Mmdilute-Va          ISARGFPSRWTYQEFFSRYRVLMKQKDVLG-----------DRKQTCKNVLEKLILDK-D 669
Rnmyo5-Va            ISAAGFPSRWTYQEFFSRYRVLMKQKDVLG-----------DRKQTCQNVLEKLILDK-D 669
Rnmyr6-Vb            ISAAGYPSRWTYHDFFNRYRVLMKKRELANTT---------DKKNICKSVLESLIKDP-D 669
HsMYO5A              ISAAGFPSRWTYQEFFSRYRVLMKQKDVLS-----------DRKQTCKNVLEKLILDK-D 669
HsMYO5C              ISAQSYPSRWTYIEFYSRYGILMTKQELSFS----------DKKEVCKVVLHRLIQDS-N 661
SpMyoVI              LMQGGFPSRTQFLDLYNMYKQYMPPELVRLDP----------RLFCKALFHALGLDE-N 687
Cehum3-VI            LMQEGFPSRTSFADLYAMYEKNLPPSLARLDP----------RLFSKCLFHALGLDQ-N 685
Dmjar-VI             LMEHGYPSRVLFADLYSMYKSVLPPELVSLPA----------RTFCEAMFQSLNLSA-K 684
Ms-VIA               LMQGGFPSRAPFHELYNMYQKYMPPKPTRLDP----------CLFCKALFKALGLNE-N 711
Ms-VIB               LMQGGFPSRASFHELYNMYKQYMPDKLTRLNP----------RLFCKALFKALGLND-S 710
GgMyoVI              LMQGGFPSRASFHELYNMYKKYLPEKLARLDP----------RLFCKALFKALGLNE-I 689
MmWaltzerVI          LMQGGFPSRASFHELYNMYKKYMPEKLPRLDP----------RLFCKPLFKALGLNE-V 692
HsMYO6               LMQGGYPSRASFHELYNMYKKYMPDKLARLDP----------RLFCKALFKALGLNE-N 689
Cehum6-VII           IRRSGYPIRHDYYPFVFRYRVLVSSIQG-PVNR-------IDLHDAAKKICHMILGTN-A 645
Dmck-VIIA            IRRAGYPIRHGFREFVERYRFLIPGVPP-AHR--------TDCQAATSRICAVVLGK--S 645
Dm28B-VIIB           IRRAGYPIRHAYRAFVERYRLLVPPVGP-LEQ--------CDCRKLARQICEVALPAD-S 646
DrMyo-VIIA           IRRAGYPIRYTFVEFVDRYRVLMPGVKP-AYKQ-------EDLRGTCQRIAEAVLGRD-D 651
Mmshaker-VIIA        IRHAGYPIRYSFVEFGERYRVLLPGVKP-AYKQ-------GDLRGTCQRMAEAVLGTH-D 651
HsUsherIb-VIIA       IRRAGYPIRYSFVEFVERYRVLLPGVKP-AYKQ-------GDLRGTCQRMAEAVLGTH-D 651
Dm10A-XV             IRQRGYPVRLRFQHFVERYRHLLPSPLARGTPY-------RELCRALLEAMPRTGVEG-P 660
Mm-MYO15             IRKEGFPVRLPFQVFIDRYRCLVALKLN-VPAD-------GDMCVSLLSRLCT--VTP-D 652
Hs-MYO15             IRKEGFPVRLPFQGFIDRYCCLVALKHD-LPAN-------GDMCVSVLSRLCK--VMP-N 652
DmMhcl-PDZ           LHRLGFPESVPLLEFVRRFGLLAGDLASNKDVSVEQILAVNELDVASYRIGPSQMWSDKS 745
Hs-PDZ               MYRQGYPDHMVFSEFRRRFDVLAPHLTKKHGR---NYIVVDERRAVEELFTLECLDLEKS 727
Mm-PDZ               MYRQGYPDHMVFSEFRRRFDVLAPHLTKKHGR---NYIVVDEKRAVEEL--LESLDLEKS 751
Dm29D                LMASGFPHRMRFKQFNARYRMLAPFRLLRRSED-KALEDCQLILKYAMEQPPVLDGSVTL 719
CONSENSUS            I-RAGFP-R--F-EF--RY-IL--------------------------LL---L-----
```




```
Ac-IB                QWQLGK----------------------------------------------SKVFIRY 655
Ac-IC                EYQKGK----------------------------------------------TKIFIRQ 646
ScMyo3-IA            EFQLGV----------------------------------------------TSVFIKT 666
ScMyo5-IB            EYQLGV----------------------------------------------TSVFIKT 666
Cehum5-IA            DCVQGR----------------------------------------------TKIFIRS 676
CehumI-IB            QYQMGK----------------------------------------------TKIFVKN 662
Dm31D-IA             DVKYGH----------------------------------------------TKIFIRS 670
Dm61F-IB             KYRVGE----------------------------------------------TKLFIRW 670
Dm95E                EFTIGT----------------------------------------------KNVFVRS 645
Rc-IB                EYKMGR----------------------------------------------TKIFIRF 694
Gg-Bb                ELAYGH----------------------------------------------TKIFIRS 695
Gg-IB                QYQMGR----------------------------------------------SKVFVKN 660
Rnmyr1-IA            EYSFGR----------------------------------------------SKIFIRN 673
Rnmyr2-IB            EYKMGR----------------------------------------------TKIFIRF 672
Rnmyr3-IC            QFQLGR----------------------------------------------SKVFIKA 660
Rnmyr4-ID            DVAYGK----------------------------------------------TKIFIRT 673
ScMYO1-IIA           VYKIGI----------------------------------------------LTVFQKL 702
ScMYS1-IIB           VYKIGN----------------------------------------------TKLFFKA 704
Dd-nmII              QYRFGI----------------------------------------------TKIFFRA 661
Cemyo3-IIA           EFKVGE----------------------------------------------TKIFFKA 693
Cemyo4-IIB           MFRIGL----------------------------------------------TKVFFKA 694
Cemyo2-IIC           NFRVGL----------------------------------------------TKVFFKA 701
Cemyo1-IID           QFRIGH----------------------------------------------TKVFFKA 689
Dm-mhcII             MYRIGH----------------------------------------------TKVFFRA 679
Dmzip-nmII           LYRVGQ----------------------------------------------SKIFFRA 723
Xl-nmIIB             LYRIGQ----------------------------------------------SKIFFRA 702
Gg-FSkII             QYRFGH----------------------------------------------TKVFFKA 682
Gg-nmII              LYRIGQ----------------------------------------------SKVFFRA 683
Gg-smII              LYRIGQ----------------------------------------------SKIFFRT 693
Hs-FSkEII            QYKFGH----------------------------------------------TKVFFKA 681
Hs-nmIIA             LYRIGQ----------------------------------------------SKVFFRA 683
Hs-CaA               QYKFGH----------------------------------------------TKVFFKA 683
DmNinaC-III          -WALGK----------------------------------------------TKVFLRY 690
Lp-III               -WTLGK----------------------------------------------NKVFLKY 635
HsMYO3A              -WALGK----------------------------------------------TKVFLKY 702
ScMyo2p-Va           KYQIGN----------------------------------------------TKIFFKA 699
ScMyo4p-Vb           KYQIGN----------------------------------------------TKIFFKA 694
Cehum2-V             KYAVGK----------------------------------------------TKIFLRT 685
SpMyo-Va             KYQFGK----------------------------------------------TKIFFRA 672
Dmdidum-Va           KYRFGN----------------------------------------------TQIFFRA 689
Ggp190-Va            KYQFGK----------------------------------------------TKIFFRA 683
Mmdilute-Va          KYQFGK----------------------------------------------TKIFFRA 682
Rnmyo5-Va            KYQFGK----------------------------------------------TKIFFRA 682
Rnmyr6-Vb            KFQFGR----------------------------------------------TKIFFRA 682
HsMYO5A              KYQFGK----------------------------------------------TKIFFRA 682
HsMYO5C              QYQFGK----------------------------------------------TKIFFRA 674
SpMyoVI              DYQFGL----------------------------------------------TKVFFRP 700
Cehum3-VI            DFQFGN----------------------------------------------TKVFFTA 698
Dmjar-VI             DFKFGI----------------------------------------------TKVFFRP 697
Ms-VIA               DYKFGLFT--------------------------------------------TIVFFRP 726
Ms-VIB               DFKFGLTF--------------------------------------------TRVFFRP 725
GgMyoVI              DYKFGL----------------------------------------------TKVFFRP 702
MmWaltzerVI          DYKFGL----------------------------------------------TQVFFRP 705
HsMYO6               DYKFGL----------------------------------------------TKVFFRP 702
Cehum6-VII           DYQLGK----------------------------------------------TKVFLKD 658
Dmck-VIIA            DYQLGH----------------------------------------------TKVFLKD 658
Dm28B-VIIB           DRQYGK----------------------------------------------TKLFLRD 659
DrMyo-VIIA           DWQMGK----------------------------------------------TKIFLKD 664
Mmshaker-VIIA        DWQIGK----------------------------------------------TKIFLKD 664
HsUsherIb-VIIA       DWQIGK----------------------------------------------TKIFLKD 664
Dm10A-XV             DYQLGA----------------------------------------------TRVFLRE 673
Mm-MYO15             MYRVGI----------------------------------------------SKLFLKE 665
Hs-MYO15             MYRVGV----------------------------------------------SKLFLKE 665
DmMhcl-PDZ           AVCLPLSGNASVAASFRNSAQILNLFYGEGHEDFVVPQPILGLLTIGCAIACVRIILFRS 805
Hs-PDZ               SCCMGLSR--------------------------------------------VFFRA 740
Mm-PDZ               SCCLGLSR--------------------------------------------VFFRA 764
Dm29D                AWAPGK----------------------------------------------RHVFLSE 732
CONSENSUS            --Y-IG----------------------------------------------TKVFFR
```

```
Ac-IB                PETLFHLEECLD---------- 667
Ac-IC                PESVFSLEELRD---------- 658
ScMyo3-IA            PESLFALEDMRD---------- 678
ScMyo5-IB            PETLFALEHMRD---------- 678
Cehum5-IA            PQTVFRLEELRT---------- 688
CehumI-IB            PESLFLLEETRE---------- 674
Dm31D-IA             PRTLFALEHQRN---------- 682
Dm61F-IB             PRTLFDTEDAYQ---------- 682
Dm95E                PRTVYELEQFRR---------- 657
Rc-IB                PKTLFATEDAFT---------- 706
Gg-Bb                PRTLFDLEKFTR---------- 707
Gg-IB                PESLFLLEEMRE---------- 672
Rnmyr1-IA            PRTLFQLEDLRK---------- 685
Rnmyr2-IB            PKTLFATEDSLE---------- 684
Rnmyr3-IC            PESLFLLEEMRE---------- 672
Rnmyr4-ID            PRTLFTLEELRA---------- 685
ScMYO1-IIA           EYWSDLEKQKD----------- 713
ScMYS1-IIB           GVLADLEKQKD----------- 715
Dd-nmII              GQLARIEEARE----------- 672
Cemyo3-IIA           GVLAKLEDLRD----------- 704
Cemyo4-IIB           GVLAHLEDIRD----------- 705
Cemyo2-IIC           GIVAHLEDLRD----------- 712
Cemyo1-IID           GVVAHIEDLRD----------- 700
Dm-mhcII             GVLGQMEEFRD----------- 690
Dmzip-nmII           GVLAHLEEERD----------- 734
Xl-nmIIB             GVLAHLEEERD----------- 713
Gg-FSkII             GLLGLLEEMRD----------- 693
Gg-nmII              GVLAHLEEERD----------- 694
Gg-smII              GVLAHLEEERD----------- 704
Hs-FSkEII            GLLGTLEEMRD----------- 692
Hs-nmIIA             GVLAHLEEERD----------- 694
Hs-CaA               GLLGLLEEMRD----------- 694
DmNinaC-III          YNDEFLARLYELQV-------- 704
Lp-III               YSEEYLSRIYETHI-------- 649
HsMYO3A              YHVEQLNLMRKEAIDKLILIQA 724
ScMyo2p-Va           GMLAYLEKLRS----------- 710
ScMyo4p-Vb           GMLAFLEKLRT----------- 705
Cehum2-V             GQVAVLERVRL----------- 696
SpMyo-Va             GQVAYLEKLRA----------- 683
Dmdidum-Va           GQVAFLEQVRA----------- 700
Ggp190-Va            GQVAYLEKIRA----------- 694
Mmdilute-Va          GQVAYLEKLRA----------- 693
Rnmyo5-Va            GQVAYLEKLRA----------- 693
Rnmyr6-Vb            GQVAYLEKLRA----------- 693
HsMYO5A              GQVAYLEKLRA----------- 693
HsMYO5C              GQVAYLEKLRL----------- 685
SpMyoVI              GKFAEFDQMMKSD--------- 713
Cehum3-VI            GKFAEFDQMMKQD--------- 711
Dmjar-VI             GKFVEFDRIMRSD--------- 710
Ms-VIA               GKFAEFDHIMKSD--------- 739
Ms-VIB               GKFAEFDQIMRSD--------- 738
GgMyoVI              GKFAEFDQIMKSD--------- 715
MmWaltzerVI          GKFAEFDQIMKSD--------- 718
HsMYO6               GKFAEFDQIMKSD--------- 715
Cehum6-VII           KHDLVLEQEYY----------- 669
Dmck-VIIA            AHDLFLEQERD----------- 669
Dm28B-VIIB           EDDASLELQRS----------- 670
DrMyo-VIIA           HHDMLLEIERD----------- 675
Mmshaker-VIIA        HHDMLLEVERD----------- 675
HsUsherIb-VIIA       HHDMLLEVERD----------- 675
Dm10A-XV             ALHRALESGRT----------- 684
Mm-MYO15             HLHQLLESMRE----------- 676
Hs-MYO15             HLYQLLESMRE----------- 676
DmMhcl-PDZ           GVLSELEAKRD----------- 816
Hs-PDZ               GTLARLEEQRD----------- 751
Mm-PDZ               GTLARLEEQRD----------- 775
Dm29D                GIRQHLEHLRT----------- 743
CONSENSUS            G--A-LE--R-----------
```

Appendix II *Myo95E mRNA*

Translated sequence is shown in black. Untranslated sequences (5'UTR and 3'UTR) are coloured in red. The promoter sequences are given in blue and the cutoff scores are shown above the sites. The transcription start sites are underlined and coloured in blue. The polyadenilation signals (AATAAA) and the subsequent cleavage sites (AAGA) are coloured in purple. The PCR primers used in the analysis of the gene are shown with arrows. The exon borders are indicated with vertical arrows.

Predicted promoters:

Score 0.78

```
  1 TAATACGCTA AAAACGAAGG CGTAGAAGGG CAGTCGGGAG TCCTGTTGTC ACGGATATTA
```

Score 0.86

```
 61 GAAGCCGGAC TCTATCTATC CGGAATATAT GAGTGAACAC GAGCTCGAGG CTGTTGCTCG
101 CCAGCACATG AAAATCAATA ATTGAATCAG AAAATCAAAA CACGGATTCG CCCAGATGCA
```

Myo95E mRNA sequence:

Ex1

```
  1 TGTTGCTCGC CAGCACATGA AAATCAATAA TTGAATCAGA AAATCAAAAC ACGGATTCGC
                EF1 (ST95F)
 61 CCAGATGCAT GCAGGCCCGC GATAAATTAT GCAATTGTTG GAGCCATTAA CAACCCCCCT
121 TACCCTTCCC ATCGTTCCCG TTCTTGTTTT GGCCCCGTTT GCAGTTGATG TATTCAAATT
181 CAAATGGTGA ACTCGATTTT GGCCGTTGTC GTTGTTGTCA CTACTGCCGT GCACGTTGTA
                                  Start Codon
241 ATTGGAGCTG GGAATCGCGA TGGAGCAGGA AATCGGCACC TGGGACTCGG TACTGTTGGA
          1BR                                   EF2 (46F)
301 GAACCTGTCC GAGGATAGTT TCATAAACAA CATCCACCAG CGCTATAAGC GCGATCACAT
```

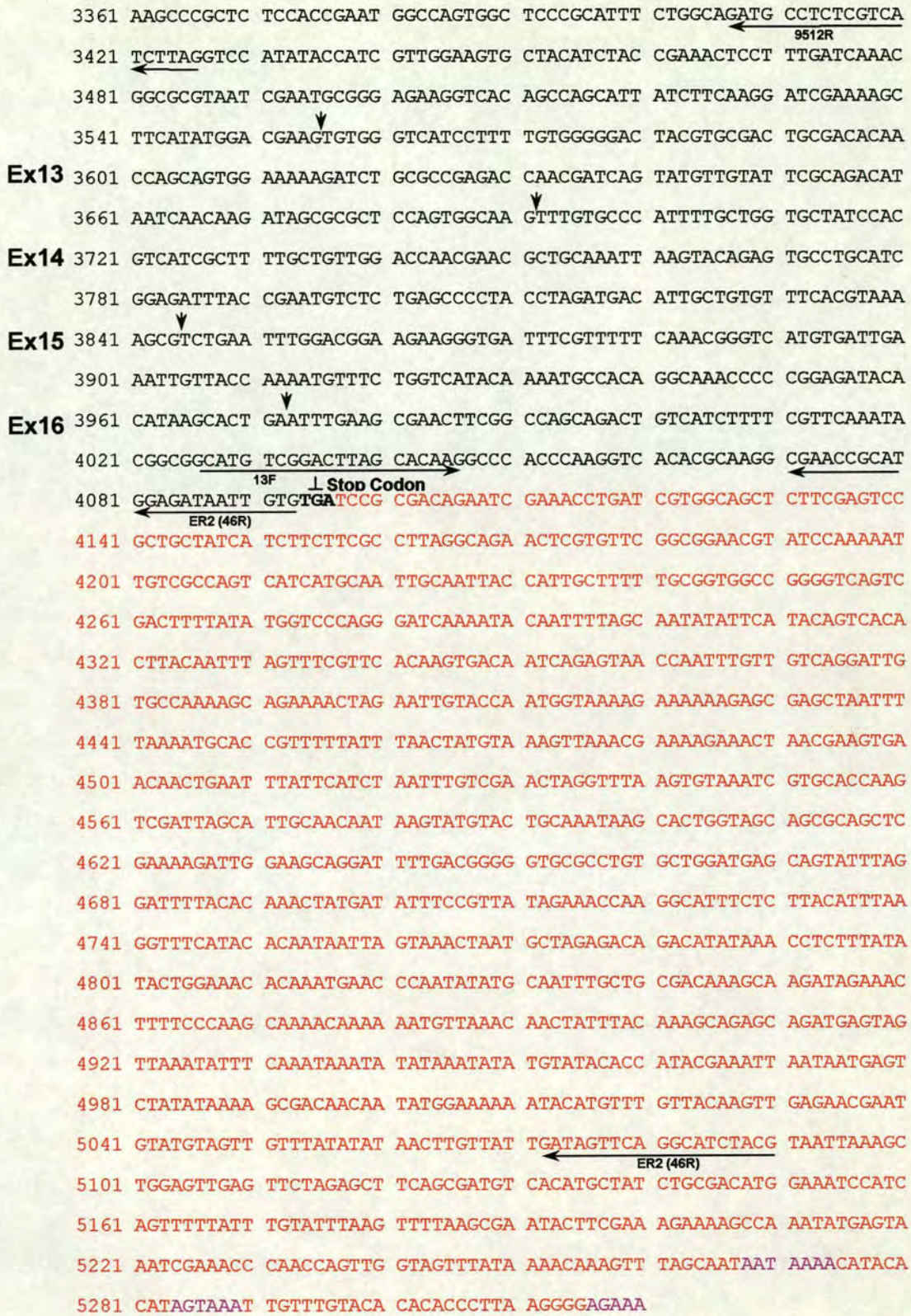
Ex2

```
361 ATATACCTAC ATTGGAACAT CTGTTGTGGC TCTGAATCCA TATCATCACA TATCCGAGCA
421 CTCTCTGGAC AATGTCCGCA ACTATGGCGA TAAGGGCATT TTCCAGCTGC CGCCCCACAT
```

Ex3

```
481 ATATGGTCTC ACAAATCTGG CTTATCAATC GCTCAAAGAT CAGAGCGAGG ATCAGTGTGT
541 TCTGCTCACC GGTGAGAGCG GAGCGGGCAA AACGGAGACT TTTAAAATGA TCGTGAACTT
601 TCTGACCCAC ATACAAGATC GCTCCCACTG CCCCCCAACA CCGAATGTTT TGCGCAAGCA
661 ATCCTCAACT AGCTCGGCCA GCGGATTGGT GATGCACGCC CACAGGCGAG CCTCCAGCAG
721 CTGCTCCGGC ACTGCCAATT TTATTATATG CAAAAACCGG GCGGAAAATC CGTCAGGCAG
781 TGTTTCACGG CGACAAAGTC CATCGCCAGG ACCATCGCAG CGATCGCGGA CGCGGGCCGA
841 GAGCATCGAG CGCCAAAGCA GGCGCCACAT GCGGGAGAAA ATTGTCGACT TTGATTTCTC
901 ACACCACAAG AGTAGCGAAA ACATCAGCGG CCTTCCTGAA TCGCACGCCC ACCACATGCA
                                  EF3 (3FOR)
961 TCCGACCAAG TCGTGCTTCA AGCACCAGCA GACCCAAGTC AGCGCCTGCA CTGCAATGCC
1021 CGCAGCAGCC AAGGGATCGC CCAAATACGC GGTACCCACC GTGTACGGCG GTTGCCGCCA
1081 GTGCGGACAC AGCAAGTGTG TCCGTGCTCA GAGCCTGGAA AAGGAGGAGC GGGATGATCT
1141 ACGAGGCAGC AACTGCCGCC TGTCTACCAT AGCCACTGCC ACCACCAATC CGGCACATCC
```


1201 GCATCGCGGT AGTTGCTCCA ACCTGATGCG CCAGCACTCG ACAGAGAGTC AGCCGGAGCG
1261 GGAGCGGGAT CGAAGCTCCC TTATGGGGTC CACGCAACGT ATATCGCTGT ACGATGCACA
1321 CAAGCTGAGT AAGGTTCTGG GCGATCTGCC GCCGCCTCCG CCTAGTTCTG TTTCGCCCAC
1381 GCCATCGGCC AGCTCTTCGC TGCACAGGCG TCATAAATCG CCCACGCAAC GAATGCGAGA
1441 GTGCGTCACC TGTGCGGATG TGTTCCTGGA GGCCATGGGC AATGCCTGCA CCCTGAAAAA
Ex4
1501 TAATAACTCT AGTCGATATG GAAAGCTATT TGATATCGAA ATTGATTTTA AGGGAGATCC
1561 AATGGGAACC CGTATAACGG ACCCTATCAT TGGAGAGCGA AATTTTCACA TCTTTTACCA
Ex5
1621 ATTACTATTA GGAGCTGATC TCCAGTTGCT AAAATCGCTC AAGCTGTATC GAAATGTGGA
Ex6
1681 AAAGTACGAG CTGCTCCGCA ACACAACTGC CATGGAGGAG GACCGCATGA ATTTTCATTA
1741 TACGAAGCGA TCCTTGGATG TGCTGGGACT TAGTTGTGAA GAGAGCAACT CCATCTTTCG
Ex7
1801 GGTCATCGCA GTGGTTCTAA AGTTGGGTAA CTTTATTTTC GTACCCATTA CTAACATTGA
4BAK
1861 TGGAACTGAG GGTTGTCAGG TGTCCAATGT ATACGAAGTT CAAGAAACAG CACAGCTTCT
Ex8
1921 TAATATGGAA GCCCAAATTC TCATAAATTG CTTAACCAGA GCGAATAGCA CGAACAGCGC
1981 TCAAGAGGAT GTGGGTTGCG AAATGGATGC ACGACAAGCA GCTACCAACC GAAACACGCT
2041 GTGTCGCACT CTTTATAGCC GTCTCTTCAC GTGGCTGGTG AACAAAATTA ATGAGTCCCT
2101 GAAGTCAACA CAGCGCGAAA AGAACCTAGC ATTGCTAGAT TTTTATGGAT TCGAAGCGCT
2161 GGACCACAAC TCATTTGAGC AATTTGCCAT TAACTACAGC GCGGAAAAGA TCCACCAGAA
Ex9
2221 TTTTGTGTTT CATGTGCTGC GTTCGGAGCA AGAGCTCTAT ATTCGCGAGG GATTGGAATG
2281 GTCTCGCATT GACTATTTCG ACAACGAGTC TATTTGCGAG TTAATAGACA AACCCAGCTA
2341 TGGTATATTG AGCTTGATTA ATGAACCCCA TTTAAATAGC AACGACGCTT TGCTTTTGCG
2401 AGTTCAGCAA TGTTGTGCGG GGCATCCCAA CTTTATGACC ACCGGCAGCA ATTCCATGTG
2461 CTTTCAGATT CGTCATTATG CAAGTGTAGT GAACTACTCA ATACATCGGT TTCTCGAAAA
2521 GAACTCCGAC ATGCTGCCGA AGTACATAAG CGCTGCCTTT TATCAGAGCA AACTTTCTTT
Ex10
ER1 (ESR)
EF4 (ESF)
2581 GGTGCAAAGC CTATTCCCCG AGGGGAATCC CCGTCGACAG GTTACCAAAA AGCCCAGCAC
2641 GTTGAGTTCG AATATCCGCA CCCAATTGCA GACGCTGCTG GCCATCGTTA AGCATCGCCG
2701 CTCCCACTAT GTGTTCTGTA TTAAGCCCAA CGAGGGCAAG CAGCCGCACC AGTTCGATAT
2761 GGCTCTAGTG CAACATCAGG TGCGCTACAT GTCGCTCATG CCGCTGGTCC ACCTGTGTCG
2821 CACTGGCCAT TGCTACCACC TGTTGCACGT TAAGTTTTTT CATCGCTATA AGTTGCTCAA
2881 CAGCCTGACG TGGCCCCACT TTCATGGCGG CAGTCAGGTA GAGGGTATCG CCCTCATAAT
2941 CCGTAACCTA CCGCTGCCCT CAGCGGAGTT CACGATCGGC ACCAAAAATG TGTTCGTGCG
3001 TAGTCCCCGC ACCGTATATG AGTTGGAACA GTTTCGCCGC CTGCGTATTA GCGAGCTGGC
P1EF
3061 CGTGCTTATT CAAACCATGT TCCGAATGTA TCACGCAAGG AAGCGCTTTC AGCGCATGCG
Ex11
3121 ACACAGCCAG ATGATCATAT CGAGTGCCTG GCGCACGTGG CGGGAATGCC GATTTGGCAT
3181 TCCCTTCACT GGTCGTAGGC ATTTGTGGAG TTTATATCGT GTCGCCCGCG AGGAGTATCG
Ex12
3241 GTCCTTGAAG TACAAACGAC AGGTGAGATG GGCCATCGAT ATTATAGGCC GCTACTACCG
3301 CCAGTGGAAG ATCAGACAGT TCCTTCTGAC AATTCCCTTG CGACTGCCAC CGAACACGCT


3361 AAGCCCGCTC TCCACCGAAT GGCCAGTGGC TCCCGCATTT CTGGCAGATG CCTCTCGTCA
9512R
3421 TCTTAGGTCC ATATACCATC GTTGGAAGTG CTACATCTAC CGAAACTCCT TTGATCAAAC
3481 GGCGCGTAAT CGAATGCGGG AGAAGGTCAC AGCCAGCATT ATCTTCAAGG ATCGAAAAGC
3541 TTCATATGGA CGAAGTGTGG GTCATCCTTT TGTGGGGGAC TACGTGCGAC TGCGACACAA
Ex13 3601 CCAGCAGTGG AAAAAGATCT GCGCCGAGAC CAACGATCAG TATGTTGTAT TCGCAGACAT
3661 AATCAACAAG ATAGCGCGCT CCAGTGGCAA GTTTGTGCCC ATTTTGCTGG TGCTATCCAC
Ex14 3721 GTCATCGCTT TTGCTGTTGG ACCAACGAAC GCTGCAAATT AAGTACAGAG TGCCTGCATC
3781 GGAGATTTAC CGAATGTCTC TGAGCCCCTA CCTAGATGAC ATTGCTGTGT TTCACGTAAA
Ex15 3841 AGCGTCTGAA TTTGGACGGA AGAAGGGTGA TTTCGTTTTT CAAACGGGTC ATGTGATTGA
3901 AATTGTTACC AAAATGTTTC TGGTCATACA AAATGCCACA GGCAAACCCC CGGAGATACA
Ex16 3961 CATAAGCACT GAATTTGAAG CGAACTTCGG CCAGCAGACT GTCATCTTTT CGTTCAAATA
4021 CGGCGGCATG TCGGACTTAG CACAAGGCCC ACCCAAGGTC ACACGCAAGG CGAACCGCAT
13F
⊥ Stop Codon
4081 GGAGATAATT GTGTGATCCG CGACAGAATC GAAACCTGAT CGTGGCAGCT CTTCGAGTCC
ER2 (46R)
4141 GCTGCTATCA TCTTCTTCGC CTTAGGCAGA ACTCGTGTTC GGCGGAACGT ATCCAAAAAT
4201 TGTCGCCAGT CATCATGCAA TTGCAATTAC CATTGCTTTT TGCGGTGGCC GGGGTCAGTC
4261 GACTTTTATA TGGTCCCAGG GATCAAAATA CAATTTTAGC AATATATTCA TACAGTCACA
4321 CTTACAATTT AGTTTCGTTC ACAAGTGACA ATCAGAGTAA CCAATTTGTT GTCAGGATTG
4381 TGCCAAAAGC AGAAAACTAG AATTGTACCA ATGGTAAAAG AAAAAAGAGC GAGCTAATTT
4441 TAAAATGCAC CGTTTTTATT TAACTATGTA AAGTTAAACG AAAAGAAACT AACGAAGTGA
4501 ACAACTGAAT TTATTCATCT AATTTGTCGA ACTAGGTTTA AGTGTAAATC GTGCACCAAG
4561 TCGATTAGCA TTGCAACAAT AAGTATGTAC TGCAAATAAG CACTGGTAGC AGCGCAGCTC
4621 GAAAAGATTG GAAGCAGGAT TTTGACGGGG GTGCGCCTGT GCTGGATGAG CAGTATTTAG
4681 GATTTTACAC AAACTATGAT ATTTCCGTTA TAGAAACCAA GGCATTTCTC TTACATTTAA
4741 GGTTTCATAC ACAATAATTA GTAAACTAAT GCTAGAGACA GACATATAAA CCTCTTTATA
4801 TACTGGAAAC ACAAATGAAC CCAATATATG CAATTTGCTG CGACAAAGCA AGATAGAAAC
4861 TTTTCCCAAG CAAAACAAAA AATGTTAAAC AACTATTTAC AAAGCAGAGC AGATGAGTAG
4921 TTAAATATTT CAAATAAATA TATAAATATA TGTATACACC ATACGAAATT AATAATGAGT
4981 CTATATAAAA GCGACAACAA TATGGAAAAA ATACATGTTT GTTACAAGTT GAGAACGAAT
5041 GTATGTAGTT GTTTATATAT AACTTGTTAT TGATAGTTCA GGCATCTACG TAATTAAAGC
ER2 (46R)
5101 TGGAGTTGAG TTCTAGAGCT TCAGCGATGT CACATGCTAT CTGCGACATG GAAATCCATC
5161 AGTTTTTATT TGTATTTAAG TTTTAAGCGA ATACTTCGAA AGAAAAGCCA AATATGAGTA
5221 AATCGAAACC CAACCAGTTG GTAGTTTATA AAACAAAGTT TAGCAATAAT AAAACATACA
5281 CATAGTAAAT TGTTTGTACA CACACCCTTA AGGGGAGAAA

Appendix III *Myo10A* mRNA

Predicted promoters:

Score 0.61

```
1 TATAATGCAA AAGACTGAGG AATTTTTATA GAAAGGCACC GTAATTAAAC GAATTTTTTT

51 TTAATTTTGT TGGGTGTTGA TAGAGAAAAA TTGTAAATTT TACTTAGTAA ATTCCCATTT
```

Score 0.4

```
101 CGTATATTTA TTGCAGCAAT GAATCAACCG GTTTGAAGAC CACATCCATA CCGCACCACT
```

Myo10A mRNA sequence:

Ex1

```
1 GCAGCAATGA ATCAACCGGT TTGAAGACCA CATCCATACC GCACCACTCC CATCAGTAGT
```

AF1 (INF)

```
61 GAGTTCCACA GATTGTCAAC ACTAAAGAGG AAGCGCCGGA GAGGAAGATC GGGAAGAGTA

121 TCGTTGCCAC ACTGAACTGA CCGGAGGATA GTGAGTCCAG AGCAGAGAAC GCTAACTTAG
```

Start Codon

```
181 CGGAGCCATG GACTGGGCCG AGGGTGATCT GGTCTGGTTC GATCCGGGCA TGGGACATCC
```

AF2 (HYST)

```
241 CATTCCCGGC GAGATTCAGG AGGTGCACCG CGCTGCCCAG GTGATTGTCG TGCAGGCGCT

301 CATCAAAGGC AAGATGCGCA TCGTACGGTT TCATTTGGAG CCGTGGCAAC GCGGCGGCGG
```

Ex2

```
361 AGGATCGGTG ACTCCCAGTC CAGCCCAATC GCCCTCTCTG GCACTCCTCT CGCCGCACGT

421 CCTGCGTCGT CAGCGTCCCG CCAGTCTCGA GGATTTCGAA GATCTTCCAA ATCTGTCGGC

481 CAAACATCCG GCTCTGCCCT CCAGTCGCCT GCGTCGCCTC AAGTGTCGAT ATCCACAAGT

541 GGTCAAGCGA CTGTCCACGC TCTGCAGCGA TGTGATAGCA CGCCATCCAC AAACATTCGC
```

Ex3

```
601 CCTGCAACCA GGAGAAGGCA GCCTGCGCGC TCGCCAGGAT CTGGGCAGCA GCGGTGTGGA

661 GGACATGACC CTGCTGGACG ATCTGCACGA GGCGTCGCTG CTGTGGAATC TTCGTCTGCG
```

510R

```
721 TTACGACAAG GGTCTAATCT ACACATTCGC CGGCAGCATC CTGATCGCGG TGAATCCGTA
```

10-AR

```
781 CAAGATGTTC CCGGATGCCT ACGGATTGGA GGTGGCCAAG CAGTATGCCG GCCGACCACT

841 TGGATCCCTG CCGCCGCATC TGTTTGCTAT CGGAGCGGCT GCACATGCCG CATTACCATC

901 GCCGCAGGTG GTAGTCATCT CCGGTGAGAG TGGCTCCGGC AAAACGGAGT CCACCAAATT

961 GGTGATGCAG TACCTGGCGG CCGTGGTGCC CGGCGGTGGA TCCGCCTCAG CAGTCATCAC

1021 CGAGCAGATC CTGGAGGCTG CTCCGCTTCT TGAGGCCTTT GGCAATGCCC GCACCGCGCG

1081 CAACGACAAT AGTTCGCGGT TTGGAAAGTA TCTGGAGGTG TACTTCAAGA GTGGCGCCAT

1141 CGTGGGCGCC AAGATTACGC AGTACCTGCT GGAGAAGTCG CGTATCGTGA CACAGGCGCC

1201 GGGCGAACGG AACTACCATG TGTTTTACGA ACTCCTAGGC GGACTATCCG AAACGGAGAG

1261 ATCCAAGTAC GGTCTGCTGG AGGCGGACAA GTACTTCTAT CTGAACCAGG GCGCCACCGA

1321 CTGTGCCAGT GGTCGAGTGG ACTGGGAGTC ACTGCAAGGG GCCATGCAGG TTCTTGGAGT

1381 CTCCGAGGGC GAGCGAGAGG GCATCGTGCG TGTCCTGGCG GCAGTTCTTC ATCTGGGAAA

1441 CGTGTACTTC CATCGCCGGC AATTGAGACA CGGCCAGGAG GGCGTGGAGG TCGGTTCCGA

1501 TGCAGAGATC AAATGGGCCG CCCACCTGCT GCACATTAGC GCGGATGGAC TGCATCGGGC

1561 ATTAACCAGT CGCACCACCG AAGCTCGAGC GGAACGATTG CACACGCCAC TTGGCATCGA
```

1621 TCAGGCATTG GATGCCCGAG ATGCCTTCGC CAAGGCTCTC TATGCGGGAC TCTTCAATTG

1681 GCTGGTCTCG CGGATCAATT CGATTGTTCA GAAGGGCGGA ACCCACGATG CCCATCGCAT

1741 CAGCATCCTG GACATCTTTG GGTTCGAGGA TTTGGCGGAG AACAGCTTCG AGCAGCTGTG

1801 CATCAACTAT GCGAATGAGA ACCTGCAGCT GTACTTCAAT AAGCATGTCT TTAAGCTGGA

1861 GCAGGCGGAG TATGCCCGCG AGCGTCTGGA GTGGACTCCG CTGGCCTGGG ACGATAATCT

1921 GCCGGTGATC CATCTACTGG CCAAGAAGCC GGTGGGCATT TGCCATCTGT TGGACGACGA

1981 GTCCAATTTC CCGCGTGCCA CCGATCTGAG TTTCCTGGAG AAGTGCCACT ACAATCATGC

2041 GCTGAGCGAG CTATACGCTC GTCCGCGGAT CGGTGCCCAG GAGTTTGGGG TGACCCACTA

2101 TGCGGGTCAG GTGTGGTACT GTGTGGACGG TTTCTTGGAC AAGAATCGCG ATGCGCTGAG

2161 GGGTGATGTC CTGGAGCTAT TGGCCTCCAG CCGGCTGCCC TTGGTGGGTG AGTTGACCAA

2221 ACAGCTTCGT GCCCAGAGAG ACGCTGGAAA AACTTTGCCA AAGGGCAGCA ATGGAAGATT

2281 CGTCACGATG AAGCCGCGAA CCCCAACGGT AGCAGCCAGA TTTGCGGACT CACTGCAGCA

2341 GTTGCTCCAA TCGATGGGCA GGTGTCATCC ATGGTTCGTG CGTTGCATCA AACCGAATCA

2401 GGAGAAGCAC GCCCTCCGCA TGGATATGCC GTGTGTGCTG CAACAACTTC GTTACTTGGG

2461 CATGCTGGAC ACCATCCAGA TCCGGCAGCG AGGCTATCCG GTGCGCTTGC GCTTCCAGCA

2521 CTTTGTGGAG CGTTATCGTC ACCTGCTGCC ATCGCCTCTG GCCAGAGGTA CACCCTACCG

2581 CGAGCTGTGC CGAGCTCTGC TGGAGGCCAT GCCGCGAACT GGTGTGGAAG GTCCGGATTA

2641 TCAATTGGGA GCTACGCGGG TTTTCCTGCG CGAGGCATTG CATCGGGCTC TGGAAAGTGG

2701 ACGAACCGAG CGTCTGCGCC GAGCGGCGGT CAGTGTCCAG AGGCATGTGC GTGGTATGCT

AF3 (P15F)

2761 TGTCCGCCGG CAATTGGCGC GTCGCCAGGC GGCCGCCACC CGTCTCCAGG CCCGTTGGCG

2821 AGGTCAGAGG GCACAGCAAC GGTACGAGCG TCTCCGCAAG GGAGCACTCA CTGCCCAGCG

2881 ATTGTGGCGA GGACGCCAGG CGAGGAGGAG AGTCCAACAA CTGCGATCCG ATCACCGACG

2941 CAGGCAGGAG GCCAGGGAGG CAGCTCAGCG AGCTCGCGAG GCACGGGAGG CCAAGCAGGC

3001 CGTGCTGGAG AGAAGTCAAC TTAGCTACCT GGACATACCA GCCGAATTGG CATTCATCTA

3061 CTCCAAACTG CAGGGCTGGT CACCGCCGCA CGGTGATCGT CATCTGGTCA AGGTATTGGG

3121 CACAGTTCCC GGTCCACCGT CGTCGGCAGT TCAGCTGCCC GAGGATTTGG GTCAGTTTTC

3181 GTTTGGAAAA TTCAGCAGCG TGTACTGCAA TGGCTTGAGA TTGCAGCCGC GGAGGGAACC

3241 CATCACTGCT CCCCTGCTAA CAAGAGCTGC CTCGAGGGAT CAGGACTTCC AGGATGCATT

3301 GGCCGTGTTC AAGCTTATCC TTCGCTGGAG TAACGATAAA GCATTGGAAG GCGCCAAGGA

3361 AAAGCTATTG GCCGACTATA TTGTGCACAA GGCTTTGAGC TCCAGAGGAT TGCGGGATGA

3421 GATTCTGGTG CAGTTGTGCA ACCAAGTTCA TGGCCTGCCG CCCAATTCCG GAGAGGCAAC

3481 TCGACTGTGG CAACTTTTGG GACAGTGTCT GTGCTGCTTC CAGCCCAGTG CTGCCTTCAG

3541 CAAATATCTG ATGAGATTCG TGGACGATGA GGCTCCAGAG TCACTGCGTC CACTACTGCT

3601 GCGCCAGTTG CTACGCCAGC AGGGCGGTGG AACGTCGAGC GGCGCAGTGG GCGCTGGCGC

3661 CTGTCGGAGC TTCGTGCCCG CCTGGCTGGA GTGGCGTGCC TGGACCAGAG GATGCGACAT

AF4 (MYTH4)

3721 GGCACTACCG CTAACTCTGC CGGACGAGGC CAGTCAAACG GTCGCCGTCG ATTCGTGGAC

3781 CAGTTGCGAG GAGGCCGCAG CTTTAGCTGT TTCATCTTTG GGCGTGGCCA GTCGTGGATG

3841 GACCCTGGTG CTGGACGATG GCCAACAGCT CACCGATAGC TGTGGACTAG ACTACGTGAT

3901 GGATTTGATT GCAGAGAAGG AGCTGTGCCC CGCATTTCCT GCTCCCAGGA GTGATCTATT

3961 GCGTTCTGGC GCCAAGTTTG CAAGAACCAC TCTGCCGGAT GCTGTTAAAC GTCCAGCGGT

4021 TCCGCCACCA GCTCCACCCA CTTCCAGTGG CAAGGAGGAT GTGCCGCGGG AACGAAGAAG

4081 CAGCCGGGAG CTACTCTCCA GAAGCTCGGC GCTCAATGAA CGCTATTTCG AACGTGAGCC

4141 CAGTCCGGGA CCGGGTCAGG GATCGAGTAC GGGTCAGGCG AAGTCGCGAT CCAAGTCACT

4201 GGATGACCTG CTCGCCGGTG ATATTGTTCC CGTGCCAACG GATTGCGATT CCCAGGAGCC

4261 GTTGCACACA CTGGGTCTGT CGGAGAGTCG CCTGAACGAT CGCTATCACT CGGCGGAGCG

4321 GCTGGCTCCA GTCGGCAAGG ACACGGGTCC TCGTTACCAG AAGTCCCAGC ACGCAGGACG

4381 CCGATCCCAT GCTGCCTCCC ATGGTTCGCA TTCCAGTAAA TATGCGGACA AGGCCGAGTA

4441 TGCCACCAGG TCCTCGGCCA TGTCCGATAC GAGTGAGGCT CCATCCTTGG CTTCCCATGT

4501 GAGGAGGGTA AGGGTCCCAT CGCAGGCATC AGATGTGGAT CAGTTCCTGG ACGATCTATT

4561 TAGTCCAGTG CTGGATGGTT CCTTGGATGA GCTATCTGAT GCCAGATCCT TGGCAGCGAG

4621 CATCCGAGGA GGCAGCTATG AACAAGAAAA TGAAGCGGAA TCGGAAATCG ATGACTTGGA

4681 TGACTACATC AACGATATAT TCCAACCGAT ACCACTCGTC CAGAACTTGG AGAAGCTGAC

4741 CAGCAAGGAT CAGTTGGCGG CTATCATCAA GGGTGGAGGA GTGGCCAGTA AACCCGAGGA

4801 GAGGGATGAG GAATCGGAAA TCGAGGATCT AGATGACTAC ATTAATGAGC TCTTTGAACC

4861 TATTCCAGTG GCCGAGGGAT TAGACAAACT GACCAGCAAG GAGCACCTAG CTGTTAGCAT

4921 CCGTGGAGGC GGCAGCACGG ACAGTAATGG AGCCGATCCG CTGCTCCACC AACTGATGCA

4981 GCTGCCCGGC GAGACGGAGT CAGGTCCAGC ACTGTATCAG CAGCAGGTGC AGCGCGCCTT

5041 CCTGCAGTCT GCAATGGCCC AGAACCTGCA GATCCAGCAG CAACTGCTCG CCCAGAATCA

5101 AGCTCTGCAG ACGCTGCTTA GTCAGCAGGC GGCAGCGGCG GCGGCCAATA ACTCCACTTC

5161 GCCGCCACCA GCTCCTCCGA TCCTGAGTAG CCTTTCTATA TCGCCACCGC CACCGCAATC

5221 CCCAATGCGC ATGAAGGCCA CCCGGAGTAG TTTGGTGGAT CCCTCGGAGA CGCGACACTT

Ex4 5281 CCTGGATCCC TACGGCAGGG CCAAGACCGT GCGTATCGGC AAGTGGCGCT GGCCACCGCC

5341 CCAGGATGAG CCCCAGTTCC AGACCGAAGA GGATTTCTTC GCCTTCAAGA TGCGCCAGCA

5401 CCAGCGCAAG ACGACGCCAC AGGCTCAGCA TCATCAGATG ACCAGCAATG GAGGTATGAT

5461 GGAGTCGGGT CCTGGAGCAC GTGGTGCCTC GGCCACGGCC ATCGAGTGGG AGGAGTTCGA

5521 GATCGAGAGT CCCACACCGC CGCCCATGAG TGGTCAGCTA ATGAGGAGTA GCATCCGACT

5581 GGAGACAACC ACAACGACAA CCACATCGGC GGTGAATAAC AGGGACAGAG ACCGCGAAAC

5641 CCAGGATGTG GGACAACCGA TGCAGAGTGT GGTCACCACA AAGCTGGCCA AGAAGAGTTT

5701 CGAGATCGGT GCGGATCGTC CGCCGCCGGG CAGTGTGGGT AAACTTAAGC TCAGTTCCGA

5761 GATGCGACAG CGGTTGGAGC AAGTAACTGC TGGCCATTCG GTTAGATCAA CCGTTTCCAC

5821 CAAGTCGGAG CAAAGAGCTC CGGCCAAACT GGAGGACACA CGGAAGCTTA TGTTGCAGCA

5881 GCAGCTGGGT GGACTCTTTG CCAGCGTGTC TGGTGGAAAT TCCGGCCCTG GTGGAGTAGT

5941 CGATTCCCAT GCAACGGTGC GCACACAGAT CGAGCGGATG GAGGGCAAAC TATCGCCACC

6001 ACCAGCTCCG CCGTCGGGCG GCTGGCCTGG AGTGCTACTG CCACCGGCTC CAAGTGTGCC

6061 GGCTCCGCCG CCACCCATCA GACCACCCAG CATGGCGCCA CCTGCCCCGC CGCCAGCTCC

6121 CCAGAGTCCG CCAACGGCTA GGAGTCCCGA ACCGGAGCCC GACTATCGCA CCTCGAGCAG

6181 TCAGGTGGTC AAGGAGCATG TGCCCGCCTT TATTCAGCGC CAGGAGCGCG ACACATTCGG

6241 TGCAGTGCGT CAGCAGCAGA TGATCAGCAG TCACCATCTG CATCTGGAGG ACCACTCAGC

6301 CTCCTCCTCA CCGGCGGTGG CCGCCTGGGA GCAGGCGGAA CGGGAAAGAT CTCGCAGCCG

6361 AAGCAGGAGT CGCGATCGTG AGGATTACTC GGAGTCGGTG TGGGATCGGG CGGAAGTGGA

6421 GGGACCCGCC TCCGGAAGTG GCAGTGAAAA GGAGCGCGAG AAGCGGGAGC GTGAACGCGA

6481 GCGTCTCTAC GAGATTCGCC AAGTGGAACG GGAGCGGGAA AGCCATAAGG TCTACCAGCC

6541 GGCGCCGCCG CGTGTCATCC AGGCTAGCAT GGACACCACC GGTGGCCATC GCCGGGAGGA

6601 CCGCACCGGC GGATTGGCCA CCTTCCGCAC CCACATGGCC CAGAAGTACG AACACGAGCG

6661 GAAGCGCAAG AGCTCGGCCA GCTCCGGAAT GCGCGAGGAG CTGGACTCGA TGCACATGAC

6721 CACGCCGCCG CCCGTTATTG TGCCGGCTCC GGTGCCACCG CCTTCGTCGA CCAGCCCGGG

6781 ATTGGGTGTG GTGTCTGGCG TTCTGATGGG ATCGGGTTCC GGGGGCAGCA CTGCATGCCT

6841 GACCTACAAC CGGGTGCCGT GGAAGTTGCG CGTGCGGAAG GAGGTCTTCC AGCCACACGA

AR2 (FERM-10) ← → FERM-12R

6901 ACCCATCGGT CCACCGGTGG CGCTGGACTT GCTCTTTGCC CAGGTGCTGG GCGACGTTTT

6961 CGGGGTGACT CCCTGCCTGC GAATTACGCC GCAGGAGAAG AGCTCCGCCC TGAACATGCT

7021 GCACGGGCAT GGCGTGAGTG TGGACACGCT GTCCAATCGA AACGGATCGG GAGCCGGCGG

7081 TGGAGGAAAT GGAGGACAGG TGCGTGCCCT GGTCAAGCGG CATCTGGTGG ATATGGCTCG

7141 CGATTGGCCG CTGTATTTCG CTCGACTGTT CGCCGTACAA GGAGCGCCGC TCTATCCGGA

7201 TGTTAGCATC ATGGGCGTCT CGCACAGTGG ACTCTACCTG GCTCGTCGAG ATGCCGACTA

7261 TCTGATCGTG GTGCAGGCCA TCTCCTTTGG CGAAATCCAG AGTGCAGTCA CCCTGCCGCG

7321 TCCTGCTGCA CTGCAGCTCA ATCTGAGGAA TGGAAAGCAT CTGGCACTCC ATGCAGCCCG

7381 CGCCGCCGCC ATCCAGTCGA TGGTCACCAC TTTTGTCCAG GAGTTCCGCA AGTCACAATC

AR3

7441 GAAGGCCTCC ACTTTGTCGT CGGGTGCCCG AGCCGCCGCC CAGACATTGA ATGTGCCGCT

7501 GGAGCGCCTG GAATCCCGCC AGGCGCACGC CCAGCGCAAC GAGCACGGCG TGGATGGCGG

7561 ATCACGACAG AATGAGCAGC TGCAGCAGTC GGAACTGCAT CATCACCTGC AGCAACAGCA

7621 GCAGCAGCAG CAGCAGCAGC AGCACCATCA GCAGCAGCAG CAACAGCAGC AGTTGGAGGA

7681 CGCTATGGAG GAGCAGCACA TGGCCACCGA TCACCAGCAG CAGCAGCAGC AGCAGTTGGG

7741 TCAGCAGCAG GGTCAACAGC AGCGCTTCCT CAAGCAGCAG AGCTACCTGC ACTCCGCCCG

7801 AAAATCGAAT GCTGGCCAAC AGCCCTCCTC GCTGACCAAC GGACAGGTGC ACCACCAGGA

7861 GCAGATGCTG CAGCAGCAAC AGCAGCAATC ACAGTCCCTA CCGCATCACG ACCTGGATGC

7921 CAACTATCTG CAGGACGAGA GTAATGGTGG CACTCCACCG TCGGTCACCA AGTACTCGCT

7881 GCTCCAGTTC GCCATGCAGC ACTTCCGCAA TGACCAACTG AGGGACGCTG ACCGCCACCA

Ex5 8041 CGAGCGGCAC CAGTCCGCCG CGAACCGCTC CTATGCGGAA CTGGTCAAGT GGCAGGGCCA

8101 CGCCATACGC CTGCCGCTCC TCCGACTTCC AAATGATTTG GCGCCACTGG CGCTCGAGTG

8161 CTTCGACTGC ATCCTGCGCT ACTGCGGTGA TATACCGCTG GATCCCGACC TCACCGAGGT

8221 CAAGTGCGTC TACACTGTGC TCATG

Appendix IV *Mhcl* mRNA

Predicted promoters:

Score 0.39

```
1 AAATCACATA AATACCGTAC TTTAAACTAT TCAGTGAATT AAAATTGTGT TACTGATTTT
```

Score.0.43

```
61 AAAAATGTAA GTTGTAATTT CAAGTTTGGA AATGAACGGT ATGCCCTGCA CGTTTGTCTA
121 GATCCACCCC TGGATAGAGC GCAGCAAAGA GAACAGTGCA AATCGAATTA GTCCAGACTT
```

Mhcl mRNA sequence:

Start Codon

Ex1

```
1 ATGTTCAACT TTATGAAGAA GAGCGCTGGG GATGAGGACA AGGAGCGTCG CAAGCGCGAA
```

MF1 (EVE1)

```
61 AAAAAACTGC GCAAAGAGGC TAAGGGCACC TCTGGTGGCC CGGGAATGGG AGGCGTGACT
121 GGCGTCACCG GCAGCATGAG CACCGACGAG CTGCTCCGCC TGGACGAGGT GCGCCGATCG
181 CTTAAAATCC GAGGGCGACG CAAGGAGAAG GAAAAGTTGC CGAGCGGAAT TACGGCGGAC
241 TATAGCGCAG ACTTCTTTGC CGCCCTCAAT GCAGATGCAA ACGGCGGAAG CAACGGAGGA
301 GCAGTGGGAG CTGCAGATCC CGAGCAGGAT CGTGGCAATG AGGAGATTGT AGCATCTGTG
361 ATGACGCACA TCGAGAGCCG AGCGTCCAAT GGCGCCGGCG TGAGTGTCAA CTACAGCGAG
```

Ex2

```
421 GTCACACATT CTCTCCTCTC TGGAGGGCTC AAGAATCGCT TTCTTCCACC TATACCCCCA
481 AAACCACCCA AGCGTGGCAT TCTGAAAGGC TCCCGCAGTA ATATGACCAA TGTCCACGAG
541 GAGATTAGCT TTGCGGGTAG TGCCGGAGGA GCAGGTGGCA CGCCAGAGAT GCTATTGCGC
```

NYP2

```
601 AACACCTTTC AGAATGAATC CCTTTACGAG GGAGCTGCCA GGGGAGCTGG TTCGATCGGC
661 TCGAACTCCA GTCAGCATGG AACCTCGTTC ACTTCCGTGG AGAGCCTGCG CAGCGATGGA
721 GAGCAGACAG GAGGTCAACA ACGGGCCATG ACCCTGCCGG CCAACGCCCG CTATGGAGCT
781 TTACCCCAGG GCAACAGTCA AGGATCGCAC CTCCATGTGC TCACCTCGCC GTCGCCCAGC
841 GCCGACAGTC TCACAGATAC CACAAACTCC TCTTTCGCCA CACCCCCATT TTCGCTGAGT
901 CCCGTCGGCG AGTCCCAAGG CATCGACCGT TGGGCCCGCG TCCATGCCTT CGAGGATGTG
961 GAGCTGCCGC TGCCGCCAGT GCAGCTAGTA AAACTTCCCC CACCAAGGCA GCTAGTGATC
```

MF2 (PDZ-1new)

```
1021 AGGCGACAGA AGTCTCCGCG CCAGGACTTC GGCTTTAGCC TGCGCAAGGC CATCTGTTTG
1081 GACAGAACAG AGTCCCTGAC TTCGCCCATT TTCAGGCCGG TTATATTCGC CGAGCCAGGT
1141 GCTGGTGGCG GAGCCACAGG TCTTTTGCCT GGCGATCGCC TAATCAAGGT CAACGGCACG
1201 CCGGTTGGGG AACTGCCGCG GGAGATCATA ATTGAGATGA TACGGAACAG CGGAGAGGCT
```

PDZ.2

```
1261 GTCACCGTGG AGGTCCAACC AGTCGCTGAG TTGGTGGAGC TGAGCAAGCG ATGCATGGCC
```

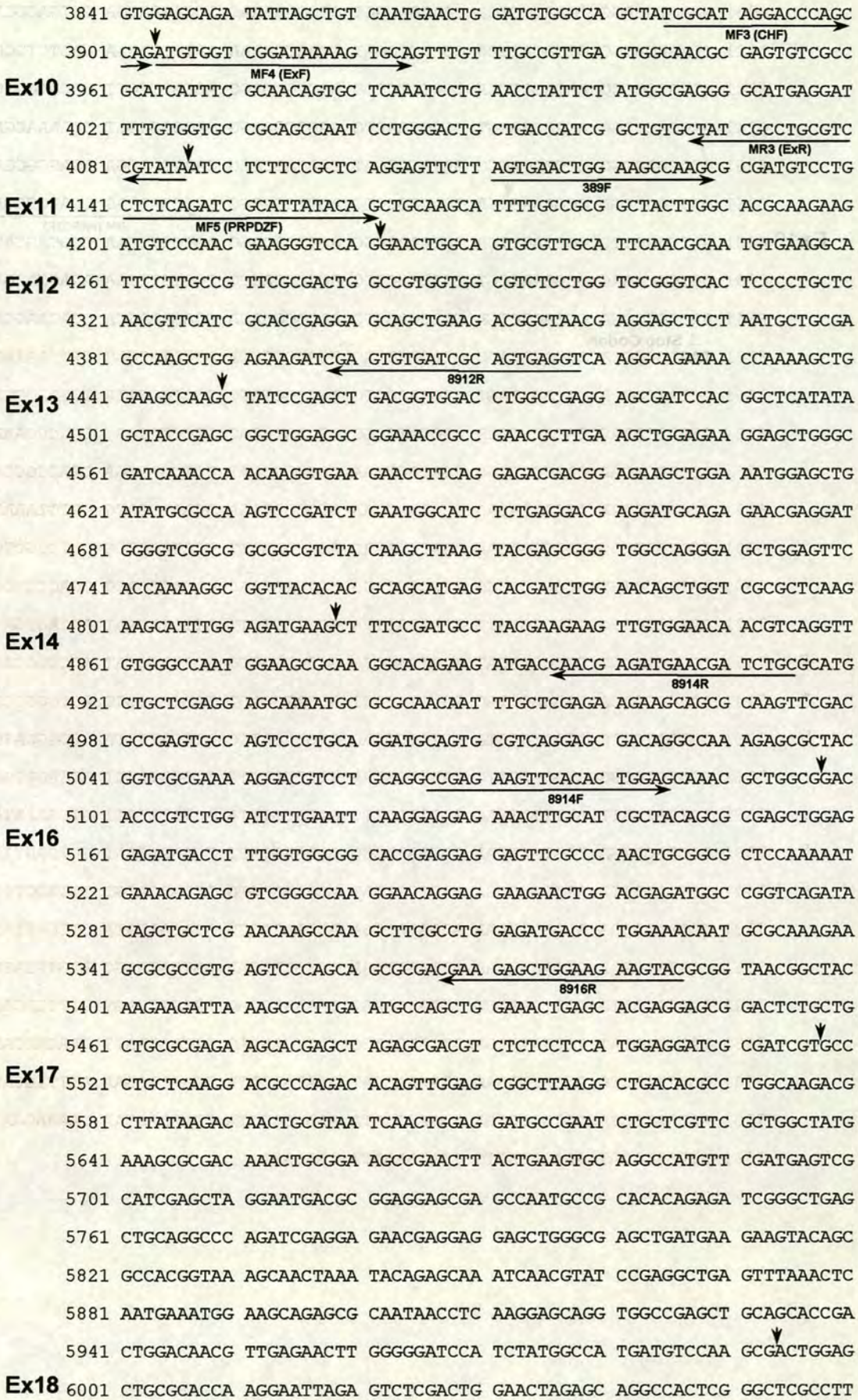
Ex3

```
1321 CCCAGCACAG CCACCGTGGA GGAGATCGAT CATTCCATTA CCAACGGCAA CTGCAATACC
1381 CTGCGACGCA GTGCCAGCAA GCGCTTCAAG CGTCAGAGCC GCCATGAGAA TGGCAATGGC
```

Ex4

```
1441 GGAGGAGAAG AGGGAGGAGC CAAAGGAGTG GGCCACGATC TGGAGGCGGA TGTTGCGGAT
1501 GCCAGTCAGC CGGAGCGCGT ATGGCTGGTG CACCGCGGTG GTTTCACTGC CGCCATCCGT
1561 CTGCCGACTG CGTCTTCTGG CCGGGATGAG GAGAACAAGC TAAGCCTGCG CCTCCTGCAT
```


1621 AATGGGGAGC AGCTGACCGT CGACGAGGAC GATGTCGAGA AGCAGAACAG CCCCGCCCTG
1681 GACCTTGTCG AGGACATCTG CGAGCTGAAG TATCTGAACG AGGCGTCCGT GCTCCATTGC
MR1 (589R)
1741 CTGCGTCAGC GATACGCCAG CAATCTCATC CACACCAAGG CGGGGCCCAC GCTGCTCGTG
1801 GTCAATCCCA TGGCACCGCT GTCCCTGTAC TCCGAGAAGG TTGTCTCCAT GTTCCGCGGC
MR2 (89R)
Ex5
1861 TGCAAGACGG AGGATATGCC GCCCCACGTC TATTCACTGG CCCAGACCGC CTACAGGAGC
1921 CTGGTCGAGA CGCGGCGCGA CCAGAGCTTG ATATTCATGG GCCGATCTGG TTCTGGCAAA
1981 TCCACTAGCT TCAAGCATGC ACTCAACTAC CTGGCGTTGG CAGCTGGTGC CTATAATAAT
Ex6
2041 TTTATAAACG CGGAGAAGGT CAATGCCCTG TGCACCATTC TGGAGGCTTT TGGAAATACG
2101 AAAACTTGCC TAAACTCGAA TGCCACGCGG ATGACGCAGC TGTTAAGCCT GGATTTCGAT
Ex7
2161 CAAACCGGCC AGATTGCATC TGCCTCATTG CAAGTTCTCC TCCCAGAAAG ACAACGAGCT
2221 GGCCGACGAC TTGGTCATGA ACATAGCTTC CACATCATGA CGAGACTTTT GGCTGGCGCT
2281 GCTGGCTTGC TACAGAAGGA GCTGCACCTG GAGAACATAA CCTCTGAGGA CAGCCATCCT
2341 TTTATATCGT TATCCCAAAA GCTGGAAGAT CGGCATCGAG CTGCCAATGA CTTTATGCGC
2401 ACTGTTCAGG CCTTTGAGAC GCTGAACATT GATGCGAAGG CGGTACGTGG GATCTGGAGC
2461 ATTTTGGCGG CCATTTACCA TCTGGGCATT GCTGGCGTCA CAAAACTGGG CACTGGCTCC
2521 ACAGCCCGAA CTCAGTTTGC CAATCCCACG GCCGCCCGAA AGGCCTCGGG CTTGTTGGGC
2581 GTGAATCTGG AGGATCTGTC ATCAGCTGCC TTTGGTCTTA CCCAACCAAA CGCCCCCAAT
2641 GGCGGTCTGA GTCCATCGAA ATCGCCCACC TCGGACACTG GACACGAATG GGCCTGGGAA
2701 TGTCTCGAGG CGTTGGTCAT TGGTCTATAT TCGGAGGCCT TGGCTGCAGT GGTGGCATTG
2761 ATCAATCGTC AGATCTGCAC ATCGTCCCAC ACCATAGCCT CAATTATGCT AATTGACACG
2821 CCCGGCTTCC AAAATCCAGC CAGCTGTGGC CAGCAAGTGG GTGCTACTCT AGCGGATCTG
2881 CGTCACAACT ATCTACAGGA GCGTCTACAG ATGCTCTTCC ATCACACCAC ACTGGTGGCT
2941 CCCCGCGATC GCTACGCCCA GGAGCTGGTA GAAATCGAGA TGGACCTGGC ATCGGAGTGT
3001 CATCCGGGAC CGTTGATCTC GCTTATCGAC AAGGCACCTC AGAATCATGT AGTGCGCTCG
3061 TCCCAGCGAG ATTTGCGAGA GCATGATCGT CGAGGAATGT TGTGGCTACT GGATGAGGAA
Ex8
3121 GCTATCTACC CCAACTCCAA CGATGACACA TTCCTTGAAC GTTTGTTCTC CCACTACGGT
3181 GACCGGGAAC ATCATAGCCT GCTGCGCAAA TGTGCAGGTC CGCGACAATT TGTACTCCAC
3241 CATCTGCAAG GCACCAATCC CGTTCTCTAC GCTGTCGATG GATGGGTGCG TCATAGCCGG
3301 GAGCACCCAG GTATTCGGAA TGCTGTATCC CTACTACAGG ACAGCAGTCG GGAGGAGATC
3361 AATCGCCTAT ATATCGGCTC ACTGACTCGA GGATCGGGAG CAATGGTGTT TTGTGGCTCT
3421 TTTGCTGGAC TGGAAGGCAC TCAGTCACTT CGTCGCGTGT CCAGTATTCG CCGCTCCTTC
3481 ACAACGGCCG GAGTAAAGCG AAACTCCATA ATGCTGCAGG TGAAGTTCAC AGTGGATGGA
3541 ATTATCGACA CTTTGCGCCG CACTGGAACT CACTTTGTGC ACTGCTATCT GCTGCAACAT
3601 AACGCAGGCA AACATACCAA GTACACCGCC AATGGATCGC CCAGCTCAGC TGCCGGTCAG
3661 GTGTCCAGCG AGGAGGAAAT GGTCAACGTG CCACTGTTAA GGAGTCAGCT ACGTGGTTCA
Ex9
3721 CAAGTCCTAG AAGCGGCTCG CTTGCATCGC CTTGGTTTTC CCGAATCAGT GCCATTGTTG
899F
3781 GAATTTGTCA GGCGTTTCGG GCTTTTGGCC GGGGATTTGG CTAGCAATAA GGATGTGAGC


3841 GTGGAGCAGA TATTAGCTGT CAATGAACTG GATGTGGCCA GCTATCGCAT AGGACCCAGC
MF3 (CHF)
3901 CAGATGTGGT CGGATAAAAG TGCAGTTTGT TTGCCGTTGA GTGGCAACGC GAGTGTCGCC
MF4 (ExF)
Ex10 3961 GCATCATTTC GCAACAGTGC TCAAATCCTG AACCTATTCT ATGGCGAGGG GCATGAGGAT
4021 TTTGTGGTGC CGCAGCCAAT CCTGGGACTG CTGACCATCG GCTGTGCTAT CGCCTGCGTC
MR3 (ExR)
4081 CGTATAATCC TCTTCCGCTC AGGAGTTCTT AGTGAACTGG AAGCCAAGCG CGATGTCCTG
389F
Ex11 4141 CTCTCAGATC GCATTATACA GCTGCAAGCA TTTTGCCGCG GCTACTTGGC ACGCAAGAAG
MF5 (PRPDZF)
4201 ATGTCCCAAC GAAGGGTCCA GGAACTGGCA GTGCGTTGCA TTCAACGCAA TGTGAAGGCA
Ex12 4261 TTCCTTGCCG TTCGCGACTG GCCGTGGTGG CGTCTCCTGG TGCGGGTCAC TCCCCTGCTC
4321 AACGTTCATC GCACCGAGGA GCAGCTGAAG ACGGCTAACG AGGAGCTCCT AATGCTGCGA
4381 GCCAAGCTGG AGAAGATCGA GTGTGATCGC AGTGAGGTCA AGGCAGAAAA CCAAAAGCTG
8912R
Ex13 4441 GAAGCCAAGC TATCCGAGCT GACGGTGGAC CTGGCCGAGG AGCGATCCAC GGCTCATATA
4501 GCTACCGAGC GGCTGGAGGC GGAAACCGCC GAACGCTTGA AGCTGGAGAA GGAGCTGGGC
4561 GATCAAACCA ACAAGGTGAA GAACCTTCAG GAGACGACGG AGAAGCTGGA AATGGAGCTG
4621 ATATGCGCCA AGTCCGATCT GAATGGCATC TCTGAGGACG AGGATGCAGA GAACGAGGAT
4681 GGGGTCGGCG GCGGCGTCTA CAAGCTTAAG TACGAGCGGG TGGCCAGGGA GCTGGAGTTC
4741 ACCAAAAGGC GGTTACACAC GCAGCATGAG CACGATCTGG AACAGCTGGT CGCGCTCAAG
Ex14 4801 AAGCATTTGG AGATGAAGCT TTCCGATGCC TACGAAGAAG TTGTGGAACA ACGTCAGGTT
4861 GTGGGCCAAT GGAAGCGCAA GGCACAGAAG ATGACCAACG AGATGAACGA TCTGCGCATG
8914R
4921 CTGCTCGAGG AGCAAAATGC GCGCAACAAT TTGCTCGAGA AGAAGCAGCG CAAGTTCGAC
4981 GCCGAGTGCC AGTCCCTGCA GGATGCAGTG CGTCAGGAGC GACAGGCCAA AGAGCGCTAC
5041 GGTCGCGAAA AGGACGTCCT GCAGGCCGAG AAGTTCACAC TGGAGCAAAC GCTGGCGGAC
8914F
Ex16 5101 ACCCGTCTGG ATCTTGAATT CAAGGAGGAG AAACTTGCAT CGCTACAGCG CGAGCTGGAG
5161 GAGATGACCT TTGGTGGCGG CACCGAGGAG GAGTTCGCCC AACTGCGGCG CTCCAAAAAT
5221 GAAACAGAGC GTCGGGCCAA GGAACAGGAG GAAGAACTGG ACGAGATGGC CGGTCAGATA
5281 CAGCTGCTCG AACAAGCCAA GCTTCGCCTG GAGATGACCC TGGAAACAAT GCGCAAAGAA
5341 GCGCGCCGTG AGTCCCAGCA GCGCGACGAA GAGCTGGAAG AAGTACGCGG TAACGGCTAC
8916R
5401 AAGAAGATTA AAGCCCTTGA ATGCCAGCTG GAAACTGAGC ACGAGGAGCG GACTCTGCTG
5461 CTGCGCGAGA AGCACGAGCT AGAGCGACGT CTCTCCTCCA TGGAGGATCG CGATCGTGCC
Ex17 5521 CTGCTCAAGG ACGCCCAGAC ACAGTTGGAG CGGCTTAAGG CTGACACGCC TGGCAAGACG
5581 CTTATAAGAC AACTGCGTAA TCAACTGGAG GATGCCGAAT CTGCTCGTTC GCTGGCTATG
5641 AAAGCGCGAC AAACTGCGGA AGCCGAACTT ACTGAAGTGC AGGCCATGTT CGATGAGTCG
5701 CATCGAGCTA GGAATGACGC GGAGGAGCGA GCCAATGCCG CACACAGAGA TCGGGCTGAG
5761 CTGCAGGCCC AGATCGAGGA GAACGAGGAG GAGCTGGGCG AGCTGATGAA GAAGTACAGC
5821 GCCACGGTAA AGCAACTAAA TACAGAGCAA ATCAACGTAT CCGAGGCTGA GTTTAAACTC
5881 AATGAAATGG AAGCAGAGCG CAATAACCTC AAGGAGCAGG TGGCCGAGCT GCAGCACCGA
5941 CTGGACAACG TTGAGAACTT GGGGGATCCA TCTATGGCCA TGATGTCCAA GCGACTGGAG
Ex18 6001 CTGCGCACCA AGGAATTAGA GTCTCGACTG GAACTAGAGC AGGCCACTCG GGCTCGCCTT

6061 GAAGTGCAGG TGAACCGTCA CAAAGAGGCT CTGGAAAAAC TGCAGAACGA GGTGACGCAG

6121 TCAAAGATGC GTGAGATGCA GGCCCAGGAT GTAATCAAAA AATCGCAAAA GAGTCTGCGC

6181 GACATGCGCG AAGAGTTCCA CGCTGTCTCC AGCCGCGAGC AGGAGTCGCT CACAAGGCGC

6241 AAGGACCTGG AGAAGAAGGT GGAGCAAATG GAGTCGGAGG GAGCGGCGCT TAAAAACGAT

6301 CTACGACTGG CTTTACAGCG AATAGCTGAT CTACAGCAGG CAATGGAGGA AGAGGGCGAG

6361 GAGGAGCTGA GCGAAAGTGA TGAAAGCCTC AGTTCGGTGG GCTCTATCAG TGATCTGGAA

MR4 (PRPDZR)

Ex19 6421 GATCGACTTC GGCCAGTTCA TGTAAAACGC AGCTCACAGC AGTCCTTGAA CGGCAGCATC

6481 GGCGGCGGAG GCGGAAGCGT TGTCAGCTCC ACCCGCACCG TGGTGTTCGA AAAGGACGAC

6541 AACAGCCCGA GGTACGCTGA TAACGAAGTC AACTCGAACA GCAGAAAGCA GCACAAGCAT

⊥ Stop Codon

6601 **TAA**AGGCCCA AGCACACAGC GTAGACTCAC TCACCCTCTA ATTATCTGAC TTTTAATATG

6661 CCTATCGCTA GTTAGCTAGT CTGTAAGGTC TGTTGCTATT TGCCGGCTTA TCTAATTTAA

6721 ATAATTAAAC CATACATGCA AACCATCCGC CTCAATCCAC ATCCACCAAC GACCCCAAAA

6781 ACCAACTCAA CAACCCCTGC CGACACTTAA ACGCTGCAAT TCTCGAACGA CTACCGCCAA

6841 TTCAAACAAA ATCATAAGAA CCGAAACGAA ACGCCCCAAA ACAAACGATG CGTGTAAATA

6901 TGTTAATGCG TATGTGACAA AAACCAAAAC ACACAAAATT TCCTAACTCC GCTCCGCTCT

6961 GATCCGCTGT ACCGCCCCCG CTCCGCTCCG CAACAACAAC AACAACAACC CGACCCTCCC

7021 AATCCGCGTG TGTGTGCGTG TAAACAATTA GCTTTTTCTT ATTGTAGAAT AACAGTGACG

7081 TCGCCATCGT CGCCACACAT TCATAAGCTG GCACTGGCGG CCAAAGCCAT ACCGGCCAAC

7141 AAACCGGACA CAAAACCTGC AGCGTCTACT AGGATGGAGT TAACAGTGCC AGCGGGCCAG

7201 CATAATGGTC ATAATGCCTA GCTGGCCAAG CAGTAGATCC ATCCATTGTA CGGGGCATGT

7261 CCATCTGGGG CTGCAGCTTG TAATTTGTAA TTTTTATTTT TTGGATATTT ATTAGTTGTT

7321 TATGATTATT ATTTTATTAC TGTTTAGTTC AATACGTTTA ACTATTTATT ATCGAAATAT

7381 AGCTACATGA TAATATTTAA TATAAAACCC AATAGGCATT TGAAAACAAT GCTGAGTTGT

7441 AGTTAATTTA ATTTGCGCGA CAGTCCCTAT ACAAAAAGTG ATTAACATGT TTTATCTCTC

7501 TCAAAAACAA AAAAACAAAA ACGAAACTGA ACTAACACCT ATAGAATATA GTCCATTATC

7561 CTATAGCAAA CTTAAATAAA ATTAGTTAAA CGCGTTTCAA AATCATAGGC AAGTGTATGC

7621 AAATACAATA AATTGAACCT TTAATCAAAT TAAAACCCAC AAATTCGAGT ATTTTACAAA

7681 TAACTCATAT TCTTATACGA ATCATCATCC CATTCACAAC ATCAGTAATT GAACATCATT

7741 AGCTTCATTC ATCATTAAAA TTAAAACCAT AAATAATATT TGAATACGAA ATTTTGCAGA

7801 AATAACGCCA AGTAACGGGA AAGATTCATA AAAACCTTTA TAAAAACTAA CAAAACGAAT

7861 ATAAACAAAT AAATATATTG TTATAAAACT TAAGAAA

Appendix V *Myo29D* mRNA

Predicted promoters:

Score 0.59
1 TTACATATAT GAGCTAAAAA GAAAGGAAAT ATCCCATCTC CCCGACGGGG AATTGAACCC

Score 0.28
61 CGGTCTCCCG CGTGACAGGC GGGGATACTT AACCACTATA CTATCGGAGG ATGTTGCGCG
121 CCGATTGCCA ATAGTTCTAT TTCAGATTCT TCGGTGCTCT CAGTGCAAAT TATTGTGCTA

Score 0.87
181 AGAATTCAAG GAAACATAAA TAACCTATTT TTGGAAAATT TCTTAAATAG CCCAACTTAT
301 CAAAACCACC CAAATATTTG AAGGTGTCTT TATTTGAACC TTCTAATTTA TCACACTTTC

Myo29D mRNA sequence:

Ex1
1 ATGGTGGCCA AAAACGCGGA GGTTCATCGA CGAGGCATTT TGGGTTTTGC TCCCCGGATC
61 GGATTGGTCA ATGTTGACTA CGACGATCTG GACAGGATTG ACAACTTCTA TTTGGAGACC
121 CAGAGGTATC GGGACTATTA CCGCGATCCG CACAACATTC TGCACAAGCC GGAGCGTTTT
DF1 (595F-A)
181 CGGCAACATC AGGGAAAATG CTCAATTAAA CTGGACACCA GTGTCCAGAA CTTGACCAGG
Ex2
241 CCGATCCATG CATCTTCATC CAACGAGGCA GTTCATCTGC AGCAGCGGCT CAAAAGCCTG
DF2 (SECF)
301 AGCACCGAAT TGGTGACTCT GCGAAACCGC CTGCATGTGG GTCATGGACC GGGTCAGGGT
361 CAGAGCCAGG GAAATGGCGC GCAGCCCGTC GCCCCAGCTC CCAATGCGGG TCCCAAGGCC
421 AACAACTTTG ACCTAAACGC ATCCAGTCCG AATCTCAATC TGAGTTCCAG CGGCGCAGGA
481 CTTTCAGCCA GTGCAGTTCA GCAACACACC AATGGACATC ACACGACATC GAAGAATCAC
Ex3
541 AGCTTCAGCC ACACATTGCC CGCAAACTCG GGCTCTGGTG GAGGAGGAGG CGCTGGAGGA
DF3 (TRIF)
601 GGCGCCGTTG TGTCGGCGAG GAACACGTCG ATTCCACATC CACTTCCGCA TCAGTTGGCC
661 GAGAAGCCGG GCCTGTCGCA TCAGCAGAGT GGCAGCGGAC ATGGTCAGTC CACCGGGACA
721 CTGCCACATA TGAGCGGCAT GGGAAGTATC CTGGGGCAAA ATTCCCATTC CCATGCACCG
781 GTCAACAACA ACAGCAACAA CTCCAACACG CTGCCCATGC GCACCTCCAA CTCCGGACAC
841 CTGGGCATCA ATGGCGGAGG CGTTGGCCAC CATCTCAGCC ACGCCCACAG CCAGCAATTG
901 CCATTCATCC CGCAGTCGAA GCACACGAAT CCCTGCCAGA GTGTGAAGAC TCTGCCCTTT
595F-B
961 GGATTCGGAT TCTCCGAAAG TCAGCAGAAG ATGCAACAGC AGCAGCAACA GGGCAACTCC
1021 TCGCCCAAGG ACATGCAGGA TCTGATCCAC TTGTCGGGAC CGCTCACCGA ACATGCCGTG
DR1 (529R)
1081 ATGCGAACGC TCCAGGCCCG ATTCAACGAG CGGCGGTACT TTACGAACGT GGGTCCAATC
Ex4
1141 TTGCTCTCCA TCAATCCGTA CCTGGACGTG GGCAACCCGC TGACCCTCAC CTCCACCCGC
1201 GCCATGCCGC TGGCCCCGCA GCTGCAGAAG ATCGTCCAGG AGGCAGTGCG TCAGCAGAGC
1261 GAGACGGGCT ATCCGCAGGC CATCATTCTC TCGGGAACAT CGGGCGCTGG AAAGACGGCC
1321 AATGCGATGC TGATGCTGCG CCAGCTCTTC GCCATCGCCG GCGGAGGTCC AGAGACTGAT
DR2 (595R-R)
1381 GCCTTCAAGC ACCTGGCAGC TGCATTCACA GTACTTCGAT CCCTGGGATC CGCCAAAACT
DF4 (595F-C)
1441 ACGACCAACT CGGAGTCCAG TCGCATTGGC CAGTTCATCG AGGTGCAGGT CACGGATGGG

```
1501 GCACTCTACC GCACCAAGAT ACACTGCTAC TTCCTGGACC AGACTCGCGT CATTCGTCCA
1561 CTGCCGAAGG AGAAGAACTA TCACATCTTC TATCAATTAC TGGCTGGTCT CAGCCGGGAG
1621 GAGCGCCAGA AGCTACACCT GGATGGCTAC TCGCCGGCCA ATCTGCGCTA CCTGCGAGGT
1681 GACATTGGCC AGAATGAGCA GGAGGATGCG GCGCGCTTTC AGGCATGGAA GACCTGCCTC
1741 GGCATCCTGG GCATCCCATT CCTGGACGTC GTTCGTGTCC TGGCCGCCGT TTTGCTGCTG
1801 GGAAATGTTC AGTTTATTGA TGGAGGGGGT CTCGAGGTGG ATGTGAAGGG TGAAACGGAG
Ex5 1861 CTGAACTCTG TGGCCAGTCT ACTTGGCGTT CCGCCGGCGG CCCTCTTCCG TGGATTGACC
1921 ACCCGTACCC ACAACGTTCG TGGCCAGCTG GTGAAGTCCG TGTGCGGCGA TGGCGATGCC
1981 AATATGACGC GGGACTGCCT GGCAAAGGCG CTCTACTGCC GCACGGTGGC CACTATTGTG
2041 CGGCGGGCCA ACAGCCTGAA GCGCCTGGGA TCCACGTTGG GCACCCTCAG CTCGGACTCA
2101 AACGAATCGG TGCACAACCA GGCGGATGTC GCCTCTCAGC ACGCCTCCAC CATTGGCGGT
2161 GGCAATGCGG GCTCCAAGTC GATGGCGGCA CTCAATAATG CAGTGCGCCA TGCCACCTCC
2221 GATGGCTTCA TTGGAATCCT GGACATGTTC GGTTTCGAGG AGCCCTCGCC ACATGCCCAT
2281 CTGGAGCACT TGTGCATCAA CCTGTGCGCC GAGACCATGC AGCACTTCTA CAACACGCAC
2341 ATCTTCAAGT CCTCCGTGGA ATCCTGTCGC GACGAGGGCA TCGTTTGCGA CACCGAGGTG
2401 GACTATGTGG ACAATGTGCC CTGCATCGAT CTGATATCCT CGCTGCGAAC CGGTCTGCTC
2461 AGCATGCTGG ATGCAGAGTG TTCGGTTAGG GGCACTGCCG AGAGTTATGT GACCAAGTTG
2521 AAGGTGCAGC ATCGCAGCTC CACTCGCTTG GAGACGAAAC CCACAGCCGA ACCACACGAT
2581 CCACGCATGT TTCTCATCCG TCACTTTGCC GGTCGGGTGG AGTACGATAC CACGGACTTT
2641 CTGGACACCA ACCGTGATGT GGTGCCCGAC GATCTGGTGG GTGTCTTCTA CAAGCACACC
2701 TGCAACTTTG GCTTCGCCAC CCATCTCTTT GGATCCGAGC TGAAAGCCCT GTATGCTCAG
2761 CAACAGGCTC CAAGGGGTCT GAGCTTCCGG ATCTCGCCCA CATCACACTC GGATCTGCTG
2821 AACGGGGACG AGCCGGTGTC CACGCTGACC CAAGACTTCC ACACCAGGCT AGACAATTTG
2881 CTGAGAACCC TGGTGCACGC CAGGCCGCAC TTTGTGCGCT GCATCCGAAG CAATGGAACC
                DF-5 (PHex)                    P29F
2941 GAGGCGGCTC GCAGCTTTGA TAGGGCTACG GTGGTGCGCC AGATTAGATC CTGCAGGTC
3001 CTGGAGACGG TCAATCTAAT GGCCAGTGGG TTCCCGCATC GCATGCGCTT TAAGCAGTTC
3061 AATGCCCGCT ACCGTATGCT AGCCCCATTC CGCCTCTTGC GCCGCAGCGA GGACAAGGCT
3121 CTGGAGGATT GCCAGTTGAT TCTGAAGTAC GCAATGGAGC AGCCGCCAGT GCTGGATGGC
3181 TCGGTGACCC TGGCCTGGGC ACCTGGCAAG CGTCATGTCT TCCTTAGCGA GGGCATTCGC
3241 CAGCACTTGG AGCACTTGCG CACCGAGATC AGGCACAAGA GCGCCACCCT GATGCAGGCC
3301 ACCTGGCGAG GCTGGTGGTG GCGCAAGAAG ATGGGCAATG GCGGGGCCAA GCGCTCCAAA
3361 ATCCCAGGTC TTCAACAGGC GCCACTGCCG AACAACAAGA CCACGCCGAA CTCCGCCGCA
3421 CAGAACAAGG CCGCATCCTC CACAATGGCC GCCCTTGCCG CTGTGGCAGC TGCTGCTCCC
3481 ATTTCTGGTG AGATTCCCCC ACGGGTATAC CCATTCTCCG CCCATCTAAT CCCCAAACTC
3541 TATGATGCAG TACCACGCCT CTCGGCCAAG ACAACGAGCC TGGGCATTGG AGGCACTGTG
3601 GCCAGACCAA GGCCACAGCC CATTGCTGGG ACTCCTCCGC CGGATCCGCA GGAGAAGTGC
Ex6 3661 GACCAGAAGA TCATTCAGCA GACCTGTAAC CTGTTTGGAG GAACGTCCGC CGCCAGTTCC
```

3721 GCCGTCTCGT TCCTACACGA TCACCGGGAA TTCGAAAATA AGCTATCCCC AAAACCGGGT

3781 CATGAAGATG AACTTTCCAG AGGAGGGACA GTCGGAGGCG CCGCAGCTGA AGAAGGGCGA

3841 AGCGGTGACC GTAGTGGGGG CCTCCACCGT TCGGGGTCAC CTGATGGTGG AGCACAAGGG

3901 GCAGAGCTTC CACGTTCCCT TCCAGTTCAT GACGCTCAGT **AA**ATAAGGAC TACTGGTGGA

3961 AACAAACTGG TTATGATTCC ATTCATTCCA TCTGTTCAAC TCTTCCTGAT TTTAGACATC

4021 CCATTTCTAA AAGACTGTAA ATATTCCATT TGGTAGATTA AATAACAAGC ATTCTGTCTA

4081 CGAAGCGCAT GCTAACTGAG GTGCTCTAAT TGGCCCCTAA CTTACGTAGT ATTATTATTG

4141 TCGCGTTAAT AAATAAACGC CAATCGATAG GTTGTAGCTT CAATTACATA CCCACAAGTG

4201 CCAATCAAAT AAGTTCTTCC ACCTAATTCA TAAGTATTTT ACATGTTACT GAGATTACAT

4261 TATTCGCCGT TAAGTAATTA GTTAAGTAAG TTTCAACAAC GACATTCATT TTTTATTTTT

4321 GAGTTCACTT AAGAATTGTG AATTTTATTA GACTAAGGTT CATTTTGTCA TTCAAATAAA

4381 CGAAGCATGC ATTGAAATAT TT

References

Adams, M. D., S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne, P. G. Amanatides, S. E. Scherer, P. W. Li, R. A. Hoskins, R. F. Galle, R. A. George, S. E. Lewis, S. Richards, M. Ashburner, S. N. Henderson, G. G. Sutton, J. R. Wortman, M. D. Yandell, Q. Zhang, L. X. Chen, R. C. Brandon, Y. H. Rogers, R. G. Blazej, M. Champe, B. D. Pfeiffer, K. H. Wan, C. Doyle, E. G. Baxter, G. Helt, C. R. Nelson, G. L. Gabor, J. F. Abril, A. Agbayani, H. J. An, C. Andrews-Pfannkoch, D. Baldwin, R. M. Ballew, A. Basu, J. Baxendale, L. Bayraktaroglu, E. M. Beasley, K. Y. Beeson, P. V. Benos, B. P. Berman, D. Bhandari, S. Bolshakov, D. Borkova, M. R. Botchan, J. Bouck, P. Brokstein, P. Brottier, K. C. Burtis, D. A. Busam, H. Butler, E. Cadieu, A. Center, I. Chandra, J. M. Cherry, S. Cawley, C. Dahlke, L. B. Davenport, P. Davies, B. de Pablos, A. Delcher, Z. Deng, A. D. Mays, I. Dew, S. M. Dietz, K. Dodson, L. E. Doup, M. Downes, S. Dugan-Rocha, B. C. Dunkov, P. Dunn, K. J. Durbin, C. C. Evangelista, C. Ferraz, S. Ferriera, W. Fleischmann, C. Fosler, A. E. Gabrielian, N. S. Garg, W. M. Gelbart, K. Glasser, A. Glodek, F. Gong, J. H. Gorrell, Z. Gu, P. Guan, M. Harris, N. L. Harris, D. Harvey, T. J. Heiman, J. R. Hernandez, J. Houck, D. Hostin, K. A. Houston, T. J. Howland, M. H. Wei, C. Ibegwam, et al. 2000. The genome sequence of Drosophila melanogaster. Science 287: 2185-95.

Al-Haddad, A., M. A. Shonn, B. Redlich, A. Blocker, J. K. Burkhardt, H. Yu, J. A. Hammer, 3rd, D. G. Weiss, W. Steffen, G. Griffiths, and S. A. Kuznetsov. 2001. Myosin Va bound to phagosomes binds to F-actin and delays microtubule- dependent motility. Mol Biol Cell 12: 2742-55.

Ashburner, M. 1974. Sequential gene activation by ecdysone in polytene chromosomes of Drosophila melanogaster. II. The effects of inhibitors of protein synthesis. Dev Biol 39: 141-57.

Avraham, K. B., T. Hasson, K. P. Steel, D. M. Kingsley, L. B. Russell, M. S. Mooseker, N. G. Copeland, and N. A. Jenkins. 1995. The mouse Snell's waltzer deafness gene encodes an unconventional myosin required for structural integrity of inner ear hair cells. Nat Genet 11: 369-75.

Bahler, M., and A. Rhoads. 2002. Calmodulin signaling via the IQ motif. FEBS Lett 513: 107-13.

Baker, J. P., and M. A. Titus. 1997. A family of unconventional myosins from the nematode Caenorhabditis elegans. J Mol Biol 272: 523-35.

Baker, J. P., and M. A. Titus. 1998. Myosins: matching functions with motors. Curr Opin Cell Biol 10: 80-6.

Bardwell, V. J., and R. Treisman. 1994. The POZ domain: a conserved protein-protein interaction motif. Genes Dev 8: 1664-77.

Barylko, B., D. D. Binns, and J. P. Albanesi. 2000. Regulation of the enzymatic and motor activities of myosin I. Biochim Biophys Acta 1496: 23-35.

Barylko, B., M. C. Wagner, O. Reizes, and J. P. Albanesi. 1992. Purification and characterization of a mammalian myosin I. Proc Natl Acad Sci U S A 89: 490-4.

Battelle, B. A., A. W. Andrews, B. G. Calman, J. R. Sellers, R. M. Greenberg, and W. C. Smith. 1998. A myosin III from Limulus eyes is a clock-regulated phosphoprotein. J Neurosci 18: 4548-59.

Bayer, C. A., B. Holley, and J. W. Fristrom. 1996a. A switch in broad-complex zinc-finger isoform expression is regulated posttranscriptionally during the metamorphosis of Drosophila imaginal discs. Dev Biol 177: 1-14.

Bayer, C. A., L. von Kalm, and J. W. Fristrom. 1996b. Gene Regulation in Imaginal Disc and Salivary Gland Development during Drosophila Metamorphosis, pp. 321-361. *In* L. I. Gilbert, B. G. Atkinson and J. R. Tata [eds.], Metamorphosis: Postembryonic Reprogramming of Gene Expression in Amphibian and Insect Cells. Academic Press.

Bayer, C. A., L. von Kalm, and J. W. Fristrom. 1997. Relationships between protein isoforms and genetic functions demonstrate functional redundancy at the Broad-Complex during Drosophila metamorphosis. Dev Biol 187: 267-82.

Bellen, H. J., C. Wilson, G. Gibson, U. Grossniklaus, R. K. Pearson, C. O'Kane, and W. J. Gehring. 1990. P-element-mediated enhancer detection allows rapid identification of developmentally regulated genes and cell specific markers in Drosophila. J Physiol 84: 33-41.

Belyaeva, E. S., M. G. Aizenzon, V. F. Semeshin, Kiss, II, K. Koczka, E. M. Baritcheva, T. D. Gorelova, and I. F. Zhimulev. 1980. Cytogenetic analysis of the 2B3-4--2B11 region of the X-chromosome of Drosophila melanogaster. I. Cytology of the region and mutant complementation groups. Chromosoma 81: 281-306.

Berg, C. A., and A. C. Spradling. 1991. Studies on the rate and site-specificity of P element transposition. Genetics 127: 515-24.

Berg, J. S., B. C. Powell, and R. E. Cheney. 2001. A millennial myosin census. Mol Biol Cell 12: 780-94.

Berg, J. S., B. H. Derfler, C. M. Pennisi, D. P. Corey, and R. E. Cheney. 2000. Myosin-X, a novel myosin with pleckstrin homology domains, associates with regions of dynamic actin. J Cell Sci 113 Pt 19: 3439-51.

Berger, B., D. B. Wilson, E. Wolf, T. Tonchev, M. Milla, and P. S. Kim. 1995. Predicting coiled coils by use of pairwise residue correlations. Proc Natl Acad Sci U S A 92: 8259-63.

Berleth, T., M. Burri, G. Thoma, D. Bopp, S. Richstein, G. Frigerio, M. Noll, and C. Nusslein-Volhard. 1988. The role of localization of bicoid RNA in organizing the anterior pattern of the Drosophila embryo. Embo J 7: 1749-56.

Bernstein, S. I., and R. A. Milligan. 1997. Fine tuning a molecular motor: the location of alternative domains in the Drosophila myosin head. J Mol Biol 271: 1-6.

Bertrand, E., P. Chartrand, M. Schaefer, S. M. Shenoy, R. H. Singer, and R. M. Long. 1998. Localization of ASH1 mRNA particles in living yeast. Mol Cell 2: 437-45.

Bezprozvanny, I., and A. Maximov. 2001. PDZ domains: More than just a glue. Proc Natl Acad Sci U S A 98: 787-9.

Biru, W. G. 1999. Drosophila melanogaster mRNA for myosin heavy chain, partial. Attila Jozsef University, Szeged, Hungary.

Bonafe, N., and J. R. Sellers. 1998. Molecular characterization of myosin V from Drosophila melanogaster. J Muscle Res Cell Motil 19: 129-41.

Bownes, M. 1989. The roles of juvenile hormone, ecdysone and the ovary in the control of Drosophila vitellogenesis. J. Insect. Physiol 35: 409-413.

Bownes, M. 1990. Preferential insertion of P elements into genes expressed in the germ- line of Drosophila melanogaster. Mol Gen Genet 222: 457-60.

Bownes, M., and L. Dale. 1982. A handbook of Drosophila development. Elsevier biomedical press, Amsterdam.

Boyne, J. R., H. M. Yosuf, P. Bieganowski, C. Brenner, and C. Price. 2000. Yeast myosin light chain, Mlc1p, interacts with both IQGAP and class II myosin to effect cytokinesis. J Cell Sci 113 Pt 24: 4533-43.

Brand, A. H., and N. Perrimon. 1993. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development 118: 401-15.

Brennan, C. A., M. Ashburner, and K. Moses. 1998. Ecdysone pathway is required for furrow progression in the developing Drosophila eye. Development 125: 2653-64.

Brennan, C. A., T. R. Li, M. Bender, F. Hsiung, and K. Moses. 2001. Broad-complex, but not ecdysone receptor, is required for progression of the morphogenetic furrow in the Drosophila eye. Development 128: 1-11.

Brockerhoff, S. E., R. C. Stevens, and T. N. Davis. 1994. The unconventional myosin, Myo2p, is a calmodulin target at sites of cell growth in Saccharomyces cerevisiae. J Cell Biol 124: 315-23.

Brown, J. L., and C. Wu. 1993. Repression of Drosophila pair-rule segmentation genes by ectopic expression of tramtrack. Development 117: 45-58.

Brown, S. S. 1997. Myosins in yeast. Curr Opin Cell Biol 9: 44-8.

Burke, T., G. L. Waring, E. Popodi, and P. Minoo. 1987. Characterization and sequence of follicle cell genes selectively expressed during vitelline membrane formation in Drosophila. Dev Biol 124: 441-50.

Burtis, K. C., C. S. Thummel, C. W. Jones, F. D. Karim, and D. S. Hogness. 1990. The Drosophila 74EF early puff contains E74, a complex ecdysone- inducible gene that encodes two ets-related proteins. Cell 61: 85-99.

Buszczak, M., and W. A. Segraves. 2000. Insect metamorphosis: out with the old, in with the new. Curr Biol 10: R830-3.

Calvi, B. R., M. A. Lilly, and A. C. Spradling. 1998. Cell cycle control of chorion gene amplification. Genes Dev 12: 734-44.

Castrillon, D. H., P. Gonczy, S. Alexander, R. Rawson, C. G. Eberhart, S. Viswanathan, S. DiNardo, and S. A. Wasserman. 1993. Toward a molecular genetic analysis of spermatogenesis in Drosophila melanogaster: characterization of male-sterile mutants generated by single P element mutagenesis. Genetics 135: 489-505.

Catlett, N. L., and L. S. Weisman. 1998. The terminal tail region of a yeast myosin-V mediates its attachment to vacuole membranes and sites of polarized growth. Proc Natl Acad Sci U S A 95: 14799-804.

Catlett, N. L., J. E. Duex, F. Tang, and L. S. Weisman. 2000. Two distinct regions in a yeast myosin-V tail domain are required for the movement of different cargoes. J Cell Biol 150: 513-26.

Chao, A. T., and G. M. Guild. 1986. Molecular analysis of the ecdysterone-inducible 2B5 "early' puff in Drosophila melanogaster. Embo J 5: 143-50.

Cheney, C. M., N. G. Kravit, and J. W. Verbsky. 1993a. A new myosin I gene in Drosophila. Biochem Biophys Res Commun 195: 1280-8.

Cheney, R. E., M. A. Riley, and M. S. Mooseker. 1993b. Phylogenetic analysis of the myosin superfamily. Cell Motil Cytoskeleton 24: 215-23.

Cheney, R. E., M. K. O'Shea, J. E. Heuser, M. V. Coelho, J. S. Wolenski, E. M. Espreafico, P. Forscher, R. E. Larson, and M. S. Mooseker. 1993c. Brain myosin-V is a two-headed unconventional myosin with motor activity. Cell 75: 13-23.

Chieregatti, E., A. Gartner, H. E. Stoffler, and M. Bahler. 1998. Myr 7 is a novel myosin IX-RhoGAP expressed in rat brain. J Cell Sci 111: 3597-608.

Christerson, L. B., and D. M. McKearin. 1994. orb is required for anteroposterior and dorsoventral patterning during Drosophila oogenesis. Genes Dev 8: 614-28.

Clark, K. A., and D. M. McKearin. 1996. The Drosophila stonewall gene encodes a putative transcription factor essential for germ cell development. Development 122: 937-50.

Cohen, D. L. 2001. Squid p196, a new member of the myosin-V class of motor proteins, is associated with motile axoplasmic organelles. Brain Res 890: 233-45.

Cope, M. J., J. Whisstock, I. Rayment, and J. Kendrick-Jones. 1996. Conservation within the myosin motor domain: implications for structure and function. Structure 4: 969-87.

Couderc, J. L., D. Godt, S. Zollman, J. Chen, M. Li, S. Tiong, S. E. Cramton, I. Sahut-Barnola, and F. A. Laski. 2002. The bric a brac locus consists of two paralogous genes encoding BTB/POZ domain proteins and acts as a homeotic and morphogenetic regulator of imaginal development in Drosophila. Development 129: 2419-33.

Crepieux, P., H. Kwon, N. Leclerc, W. Spencer, S. Richard, R. Lin, and J. Hiscott. 1997. I kappaB alpha physically interacts with a cytoskeleton-associated protein through its signal response domain. Mol Cell Biol 17: 7375-85.

Crispi, S., E. Giordano, P. P. D'Avino, and M. Furia. 1998. Cross-talking among Drosophila nuclear receptors at the promiscuous response element of the ng-1 and ng-2 intermolt genes. J Mol Biol 275: 561-74.

de Cuevas, M., M. A. Lilly, and A. C. Spradling. 1997. Germline cyst formation in Drosophila. Annu Rev Genet 31: 405-28.

De La Cruz, E. M., H. L. Sweeney, and E. M. Ostap. 2000. ADP inhibition of myosin V ATPase activity. Biophys J 79: 1524-9.

Delidakis, C., and F. C. Kafatos. 1989. Amplification enhancers and replication origins in the autosomal chorion gene cluster of Drosophila. Embo J 8: 891-901.

Deng, W., K. Leaper, and M. Bownes. 1999. A targeted gene silencing technique shows that Drosophila myosin VI is required for egg chamber and imaginal disc morphogenesis. J Cell Sci 112: 3677-90.

Deng, W. M. 1997. Functional characterisation of two genes expressed in subsets of follicle cells during Drosophila oogenesis, pp. 213, ICMB. The University of Edinburgh, Edinburgh.

Deng, W. M., and M. Bownes. 1997. Two signalling pathways specify localised expression of the Broad- Complex in Drosophila eggshell patterning and morphogenesis. Development 124: 4639-47.

Deng, W. M., and M. Bownes. 1998. Patterning and morphogenesis of the follicle cell epithelium during Drosophila oogenesis. Int J Dev Biol 42: 541-52.

Deng, W. M., D. Zhao, K. Rothwell, and M. Bownes. 1997. Analysis of P[gal4] insertion lines of Drosophila melanogaster as a route to identifying genes important in the follicle cells during oogenesis. Mol Hum Reprod 3: 853-62.

DiBello, P. R., D. A. Withers, C. A. Bayer, J. W. Fristrom, and G. M. Guild. 1991. The Drosophila Broad-Complex encodes a family of related proteins containing zinc fingers. Genetics 129: 385-97.

Dobens, L. L., and L. A. Raftery. 2000. Integration of epithelial patterning and morphogenesis in Drosophila ovarian follicle cells. Dev Dyn 218: 80-93.

Dose, A. C., and B. Burnside. 2002. A class III myosin expressed in the retina is a potential candidate for Bardet-Biedl syndrome. Genomics 79: 621-4.

Edwards, K. A., and D. P. Kiehart. 1996. Drosophila nonmuscle myosin II has multiple essential roles in imaginal disc and egg chamber morphogenesis. Development 122: 1499-511.

Edwards, K. A., X. J. Chang, and D. P. Kiehart. 1995. Essential light chain of Drosophila nonmuscle myosin II. J Muscle Res Cell Motil 16: 491-8.

Eggleston, W. B. 1990. P element transposition and excision in Drosophila: Interactions between elements. Wisconsin, Genetics 203.

Emery, I. F., V. Bedian, and G. M. Guild. 1994. Differential expression of Broad-Complex transcription factors may forecast tissue-specific developmental fates during Drosophila metamorphosis. Development 120: 3275-87.

Engle, L. J., and R. H. Kennett. 1994. Cloning, analysis, and chromosomal localization of myoxin (MYH12), the human homologue to the mouse dilute gene. Genomics 19: 407-16.

Espindola, F. S., E. M. Espreafico, M. V. Coelho, A. R. Martins, F. R. Costa, M. S. Mooseker, and R. E. Larson. 1992. Biochemical and immunological characterization of p190-calmodulin complex from vertebrate brain: a novel calmodulin-binding myosin. J Cell Biol 118: 359-68.

Espindola, F. S., D. M. Suter, L. B. Partata, T. Cao, J. S. Wolenski, R. E. Cheney, S. M. King, and M. S. Mooseker. 2000. The light chain composition of chicken brain myosin-Va: calmodulin, myosin-II essential light chains, and 8-kDa dynein light chain/PIN. Cell Motil Cytoskeleton 47: 269-81.

Espreafico, E. M., R. E. Cheney, M. Matteoli, A. A. Nascimento, P. V. De Camilli, R. E. Larson, and M. S. Mooseker. 1992. Primary structure and cellular localization of chicken brain myosin-V (p190), an unconventional myosin with calmodulin light chains. J Cell Biol 119: 1541-57.

Espreafico, E. M., D. E. Coling, V. Tsakraklides, K. Krogh, J. S. Wolenski, G. Kalinec, and B. Kachar. 1998. Localization of myosin-V in the centrosome. Proc Natl Acad Sci U S A 95: 8636-41.

Evans, L. L., J. Hammer, and P. C. Bridgman. 1997. Subcellular localization of myosin V in nerve growth cones and outgrowth from dilute-lethal neurons. J Cell Sci 110: 439-49.

Fedorova, S. A., S. Nokkala, and L. V. Omel'ianchuk. 2001. [Genetic screening of meiotic mutations in mosaic clones of the Drosophila melanogaster female germline]. Genetika 37: 1621-31.

Forbes, A. J., Y. Nakano, A. M. Taylor, and P. W. Ingham. 1993. Genetic analysis of hedgehog signalling in the Drosophila embryo. Dev Suppl: 115-24.

Friedberg, F., and A. R. Rhoads. 2001. Evolutionary aspects of calmodulin. IUBMB Life 51: 215-21.

Friedman, T. B., J. R. Sellers, and K. B. Avraham. 1999. Unconventional myosins and the genetics of hearing loss. Am J Med Genet 89: 147-57.

Fujiwara, M., H. Horiuchi, A. Ohta, and M. Takagi. 1997. A novel fungal gene encoding chitin synthase with a myosin motor-like domain. Biochem Biophys Res Commun 236: 75-8.

Furusawa, T., S. Ikawa, N. Yanai, and M. Obinata. 2000. Isolation of a novel PDZ-containing myosin from hematopoietic supportive bone marrow stromal cell lines. Biochem Biophys Res Commun 270: 67-75.

Geli, M. I., R. Lombardi, B. Schmelzl, and H. Riezman. 2000. An intact SH3 domain is required for myosin I-induced actin polymerization. Embo J 19: 4281-91.

Ghbeish, N., C. C. Tsai, M. Schubiger, J. Y. Zhou, R. M. Evans, and M. McKeown. 2001. The dual role of ultraspiracle, the Drosophila retinoid X receptor, in the ecdysone response. Proc Natl Acad Sci U S A 98: 3867-72.

Gilbert, S. P., and A. T. Mackey. 2000. Kinetics: a tool to study molecular motors. Methods 22: 337-54.

Godt, D., and F. A. Laski. 1995. Mechanisms of cell rearrangement and cell recruitment in Drosophila ovary morphogenesis and the requirement of bric a brac. Development 121: 173-87.

Golic, K. G. 1994. Local transposition of P elements in Drosophila melanogaster and recombination between duplicated elements using a site-specific recombinase. Genetics 137: 551-63.

Gonzalez-Reyes, A., and D. St Johnston. 1994. Role of oocyte position in establishment of anterior-posterior polarity in Drosophila. Science 266: 639-42.

Gonzalez-Reyes, A., H. Elliott, and D. St Johnston. 1995. Polarization of both major body axes in Drosophila by gurken-torpedo signalling. Nature 375: 654-8.

Gonzy, G., G. V. Pokholkova, F. Peronnet, B. Mugat, O. V. Demakova, I. V. Kotlikova, J. A. Lepesant, and I. F. Zhimulev. 2002. Isolation and characterization of novel mutations of the Broad-Complex, a key regulatory gene of ecdysone induction in Drosophila melanogaster. Insect Biochem Mol Biol 32: 121-32.

Goode, S., M. Melnick, T. B. Chou, and N. Perrimon. 1996. The neurogenic genes egghead and brainiac define a novel signaling pathway essential for epithelial morphogenesis during Drosophila oogenesis. Development 122: 3863-79.

Grad, I., A. Niedziela-Majka, M. Kochman, and A. Ozyhar. 2001. Analysis of Usp DNA binding domain targeting reveals critical determinants of the ecdysone receptor complex interaction with the response element. Eur J Biochem 268: 3751-8.

Grossniklaus, U., H. J. Bellen, C. Wilson, and W. J. Gehring. 1989. P-element-mediated enhancer detection applied to the study of oogenesis in Drosophila. Development 107: 189-200.

Guo, M., E. Bier, L. Y. Jan, and Y. N. Jan. 1995. tramtrack acts downstream of numb to specify distinct daughter cell fates during asymmetric cell divisions in the Drosophila PNS. Neuron 14: 913-25.

Haarer, B. K., A. Petzold, S. H. Lillie, and S. S. Brown. 1994. Identification of MYO4, a second class V myosin gene in yeast. J Cell Sci 107: 1055-64.

Hames, B. D., and M. Bownes. 1978. Synthesis of yolk proteins in Drosophila melanogaster. Insect. Biochem. 8: 319-328.

Hammer, J. A., 3rd, and G. Jung. 1996. The sequence of the dictyostelium myo J heavy chain gene predicts a novel, dimeric, unconventional myosin with a heavy chain molecular mass of 258 kDa. J Biol Chem 271: 7120-7.

Harrison, S. D., and A. A. Travers. 1990. The tramtrack gene encodes a Drosophila finger protein that interacts with the ftz transcriptional regulatory region and shows a novel embryonic expression pattern. Embo J 9: 207-16.

Hasson, T., and M. S. Mooseker. 1995. Molecular motors, membrane movements and physiology: emerging roles for myosins. Curr Opin Cell Biol 7: 587-94.

Hasson, T., and R. E. Cheney. 2001. Mechanisms of motor protein reversal. Curr Opin Cell Biol 13: 29-35.

Hasson, T., P. G. Gillespie, J. A. Garcia, R. B. MacDonald, Y. Zhao, A. G. Yee, M. S. Mooseker, and D. P. Corey. 1997. Unconventional myosins in inner-ear sensory epithelia. J Cell Biol 137: 1287-307.

Hasson, T., J. F. Skowron, D. J. Gilbert, K. B. Avraham, W. L. Perry, W. M. Bement, B. L. Anderson, E. H. Sherr, Z. Y. Chen, L. A. Greene, D. C. Ward, D. P. Corey, M. S. Mooseker, N. G. Copeland, and N. A. Jenkins. 1996. Mapping of unconventional myosins in mouse and human. Genomics 36: 431-9.

Hastings, G. A., and C. P. Emerson, Jr. 1991. Myosin functional domains encoded by alternative exons are expressed in specific thoracic muscles of Drosophila. J Cell Biol 114: 263-76.

Hayden, S. M., J. S. Wolenski, and M. S. Mooseker. 1990. Binding of brush border myosin I to phospholipid vesicles. J Cell Biol 111: 443-51.

Heitzler, P., D. Coulson, M. T. Saenz-Robles, M. Ashburner, J. Roote, P. Simpson, and D. Gubb. 1993. Genetic and cytogenetic analysis of the 43A-E region containing the segment polarity gene costa and the cellular polarity genes prickle and spiny-legs in Drosophila melanogaster. Genetics 135: 105-15.

Henrich, V. C., and N. E. Brown. 1995. Insect nuclear receptors: a developmental and comparative perspective. Insect Biochem Mol Biol 25: 881-97.

Henrich, V. C., A. A. Szekely, S. J. Kim, N. E. Brown, C. Antoniewski, M. A. Hayden, J. A. Lepesant, and L. I. Gilbert. 1994. Expression and function of the ultraspiracle (usp) gene during development of Drosophila melanogaster. Dev Biol 165: 38-52.

Hettmann, C., A. Herm, A. Geiter, B. Frank, E. Schwarz, T. Soldati, and D. Soldati. 2000. A dibasic motif in the tail of a class XIV apicomplexan myosin is an essential determinant of plasma membrane localization. Mol Biol Cell 11: 1385-400.

Hicks, J. L., and D. S. Williams. 1992. Distribution of the myosin I-like ninaC proteins in the Drosophila retina and ultrastructural analysis of mutant phenotypes. J Cell Sci 101: 247-54.

Hicks, J. L., W. M. Deng, A. D. Rogat, K. G. Miller, and M. Bownes. 1999. Class VI unconventional myosin is required for spermatogenesis in Drosophila. Mol Biol Cell 10: 4341-53.

Highsmith, S. 1999. Lever arm model of force generation by actin-myosin-ATP. Biochemistry 38: 9791-7.

Hodge, T., and M. J. Cope. 2000. A myosin family tree. J Cell Sci 113 Pt 19: 3353-4.

Hodgetts, R. B., W. C. Clark, S. L. O'Keefe, M. Schouls, K. Crossgrove, G. M. Guild, and L. von Kalm. 1995. Hormonal induction of Dopa decarboxylase in the epidermis of Drosophila is mediated by the Broad-Complex. Development 121: 3913-22.

Holme, R. H., and K. P. Steel. 1999. Genes involved in deafness. Curr Opin Genet Dev 9: 309-14.

Homma, K., J. Saito, R. Ikebe, and M. Ikebe. 2000. Ca(2+)-dependent regulation of the motor activity of myosin V. J Biol Chem 275: 34766-71.

Horner, M. A., T. Chen, and C. S. Thummel. 1995. Ecdysteroid regulation and DNA binding properties of Drosophila nuclear hormone receptor superfamily members. Dev Biol 168: 490-502.

Horowitz, J. A., and J. A. Hammer, 3rd. 1990. A new Acanthamoeba myosin heavy chain. Cloning of the gene and immunological identification of the polypeptide. J Biol Chem 265: 20646-52.

Huang, R. Y., and W. C. Orr. 1992. Broad-complex function during oogenesis in Drosophila melanogaster. Dev Genet 13: 277-88.

Ikonen, E., J. B. de Almeid, K. R. Fath, D. R. Burgess, K. Ashman, K. Simons, and J. L. Stow. 1997. Myosin II is associated with Golgi membranes: identification of p200 as nonmuscle myosin II on Golgi-derived vesicles. J Cell Sci 110: 2155-64.

Ikura, M. 1996. Calcium binding and conformational response in EF-hand proteins. Trends Biochem Sci 21: 14-7.

Isaac, P. G., and M. Bownes. 1982. Ovarian and fat-body vitellogenin synthesis in Drosophila melanogaster. Eur J Biochem 123: 527-34.

Jansen, R. P., C. Dowzer, C. Michaelis, M. Galova, and K. Nasmyth. 1996. Mother cell-specific HO expression in budding yeast depends on the unconventional myosin myo4p and other cytoplasmic proteins. Cell 84: 687-97.

Johnston, G. C., J. A. Prendergast, and R. A. Singer. 1991. The Saccharomyces cerevisiae MYO2 gene encodes an essential myosin for vectorial transport of vesicles. J Cell Biol 113: 539-51.

Kania, A., A. Salzberg, M. Bhat, D. D'Evelyn, Y. He, I. Kiss, and H. J. Bellen. 1995. P-element mutations affecting embryonic peripheral nervous system development in Drosophila melanogaster. Genetics 139: 1663-78.

Karess, R. E., X. J. Chang, K. A. Edwards, S. Kulkarni, I. Aguilera, and D. P. Kiehart. 1991. The regulatory light chain of nonmuscle myosin is encoded by spaghetti-squash, a gene required for cytokinesis in Drosophila. Cell 65: 1177-89.

Karim, F. D., G. M. Guild, and C. S. Thummel. 1993. The Drosophila Broad-Complex plays a key role in controlling ecdysone- regulated gene expression at the onset of metamorphosis. Development 118: 977-88.

Kashiyama, T., N. Kimura, T. Mimura, and K. Yamamoto. 2000. Cloning and characterization of a myosin from characean alga, the fastest motor protein in the world. J Biochem (Tokyo) 127: 1065-70.

Kawasaki, H., S. Nakayama, and R. H. Kretsinger. 1998. Classification and evolution of EF-hand proteins. Biometals 11: 277-95.

Kellerman, K. A., and K. G. Miller. 1992. An unconventional myosin heavy chain gene from Drosophila melanogaster. J Cell Biol 119: 823-34.

Kelley, M. R., S. Kidd, R. L. Berg, and M. W. Young. 1987. Restriction of P-element insertions at the Notch locus of Drosophila melanogaster. Mol Cell Biol 7: 1545-8.

Keyes, L. N., and A. C. Spradling. 1997. The Drosophila gene fs(2)cup interacts with otu to define a cytoplasmic pathway required for the structure and function of germ-line chromosomes. Development 124: 1419-31.

Khan, M. M., and M. Komiyama. 2001. The second EF-hand is responsible for the isoform-specific sorting of myosin essential light chain. Cell Struct Funct 26: 243-51.

Kiehart, D. P., and R. Feghali. 1986. Cytoplasmic myosin from Drosophila melanogaster. J Cell Biol 103: 1517-25.

Kiehart, D. P., R. A. Montague, J. Roote, and M. Ashburner. 1998. Evidence that crinkled, mutations in which cause numerous defects in Drosophila morphogenesis, encodes a myosin VII. Mol. Biol. Cell Suppl. 9: 388a.

Kiehart, D. P., C. G. Galbraith, K. A. Edwards, W. L. Rickoll, and R. A. Montague. 2000. Multiple forces contribute to cell sheet morphogenesis for dorsal closure in Drosophila. J Cell Biol 149: 471-90.

King, R. C. 1970. Ovarian development in Drosophila melanogaster. Academic Press, New York.

Kinkema, M., and J. Schiefelbein. 1994. A myosin from a higher plant has structural similarities to class V myosins. J Mol Biol 239: 591-7.

Kinkema, M., H. Wang, and J. Schiefelbein. 1994. Molecular analysis of the myosin gene family in Arabidopsis thaliana. Plant Mol Biol 26: 1139-53.

Kiss, I., A. H. Beaton, J. Tardiff, D. Fristrom, and J. W. Fristrom. 1988. Interactions and developmental effects of mutations in the Broad- Complex of Drosophila melanogaster. Genetics 118: 247-59.

Knowles, B. A., and L. Cooley. 1994. The specialized cytoskeleton of the Drosophila egg chamber. Trends Genet 10: 235-41.

Komitopoulou, K., M. Gans, L. H. Margaritis, K. F. C., and M. Masson. 1983. Isolation and characterization of sex-linked female sterile mutants in Drosophila melanogaster with special attention to egg-shell mutants. Genetics: 893-920.

Korn, E. D. 2000. Coevolution of head, neck, and tail domains of myosin heavy chains. Proc Natl Acad Sci U S A 97: 12559-64.

Lai, Z. C., and Y. Li. 1999. Tramtrack69 is positively and autonomously required for Drosophila photoreceptor development. Genetics 152: 299-305.

Lane, M. E., and D. Kalderon. 1995. Localization and functions of protein kinase A during Drosophila oogenesis. Mech Dev 49: 191-200.

Langford, G. M. 1995. Actin- and microtubule-dependent organelle motors: interrelationships between the two motility systems. Curr Opin Cell Biol 7: 82-8.

Lantz, V., J. S. Chang, J. I. Horabin, D. Bopp, and P. Schedl. 1994. The Drosophila orb RNA-binding protein is required for the formation of the egg chamber and establishment of polarity. Genes Dev 8: 598-613.

Lantz, V. A., and K. G. Miller. 1998. A class VI unconventional myosin is associated with a homologue of a microtubule-binding protein, cytoplasmic linker protein-170, in neurons and at the posterior pole of Drosophila embryos. J Cell Biol 140: 897-910.

Lapierre, L. A., R. Kumar, C. M. Hales, J. Navarre, S. G. Bhartur, J. O. Burnette, D. W. Provance, Jr., J. A. Mercer, M. Bahler, and J. R. Goldenring. 2001. Myosin vb is associated with plasma membrane recycling systems. Mol Biol Cell 12: 1843-57.

Larkin, M. K., K. Holder, C. Yost, E. Giniger, and H. Ruohola-Baker. 1996. Expression of constitutively active Notch arrests follicle cells at a precursor stage during Drosophila oogenesis and disrupts the anterior- posterior axis of the oocyte. Development 122: 3639-50.

Larson, R. E. 1996. Myosin-V: a class of unconventional molecular motors. Braz J Med Biol Res 29: 309-18.

Lasko, P. F. 1994. Molecular Genetics of Drosophila Oogenesis (Molecular Biology Intelligence Unit). R G Landes Co.

Lee, W. L., E. M. Ostap, H. G. Zot, and T. D. Pollard. 1999. Organization and ligand binding properties of the tail of Acanthamoeba myosin-IA. Identification of an actin-binding site in the basic (tail homology-1) domain. J Biol Chem 274: 35159-71.

Li, H. S., J. A. Porter, and C. Montell. 1998. Requirement for the NINAC kinase/myosin for stable termination of the visual cascade. J Neurosci 18: 9601-6.

Liang, Y., A. Wang, I. A. Belyantseva, D. W. Anderson, F. J. Probst, T. D. Barber, W. Miller, J. W. Touchman, L. Jin, S. L. Sullivan, J. R. Sellers, S. A. Camper, R. V. Lloyd, B. Kachar, T. B. Friedman, and R. A. Fridell. 1999. Characterization of the human and mouse unconventional myosin XV genes responsible for hereditary deafness DFNB3 and shaker 2. Genomics 61: 243-58.

Lin, H., and A. C. Spradling. 1995. Fusome asymmetry and oocyte determination in Drosophila. Dev Genet 16: 6-12.

Liu, L., J. Zhou, and T. C. Pesacreta. 2001. Maize myosins: diversity, localization, and function. Cell Motil Cytoskeleton 48: 130-48.

Liu, X., H. Brzeska, and E. D. Korn. 2000. Functional analysis of tail domains of Acanthamoeba myosin IC by characterization of truncation and deletion mutants. J Biol Chem 275: 24886-92.

Lloyd, R. V., S. Vidal, L. Jin, S. Zhang, K. Kovacs, E. Horvath, B. W. Scheithauer, E. T. Boger, R. A. Fridell, and T. B. Friedman. 2001. Myosin XVA expression in the pituitary and in other neuroendocrine tissues and tumors. Am J Pathol 159: 1375-82.

MacIver, B., A. McCormack, R. Slee, and M. Bownes. 1998. Identification of an essential gene encoding a class-V unconventional myosin in Drosophila melanogaster. Eur J Biochem 257: 529-37.

Maki, M., S. V. Narayana, and K. Hitomi. 1997. A growing family of the Ca2+-binding proteins with five EF-hand motifs. Biochem J 328: 718-20.

Mansfield, S. G., D. Y. al-Shirawi, A. S. Ketchum, E. C. Newbern, and D. P. Kiehart. 1996. Molecular organization and alternative splicing in zipper, the gene that encodes the Drosophila non-muscle myosin II heavy chain. J Mol Biol 255: 98-109.

Margolis, J., and A. Spradling. 1995. Identification and behavior of epithelial stem cells in the Drosophila ovary. Development 121: 3797-807.

McCormack, A. 1998. Characterisation of a gene encoding a transcript localised during oogenesis in Drosophila melanogaster, pp. 267, ICMB. The University of Edinburgh, Edinburgh.

McCormack, A., B. MacIver, and M. Bownes. 1998. Cloning and expression of az2, a putative zinc finger transcription factor from Drosophila melanogaster. Dev Genes Evol 208: 172-4.

McKearin, D. 1997. The Drosophila fusome, organelle biogenesis and germ cell differentiation: if you build it. Bioessays 19: 147-52.

Mercer, J. A., P. K. Seperack, M. C. Strobel, N. G. Copeland, and N. A. Jenkins. 1991. Novel myosin heavy chain encoded by murine dilute coat colour locus. Nature 349: 709-13.

Mermall, V., and K. G. Miller. 1995. The 95F unconventional myosin is required for proper organization of the Drosophila syncytial blastoderm. J Cell Biol 129: 1575-88.

Mermall, V., P. L. Post, and M. S. Mooseker. 1998. Unconventional myosins in cell movement, membrane traffic, and signal transduction. Science 279: 527-33.

Micklem, D. R., R. Dasgupta, H. Elliott, F. Gergely, C. Davidson, A. Brand, A. Gonzalez-Reyes, and D. St Johnston. 1997. The mago nashi gene is required for the polarisation of the oocyte and the formation of perpendicular axes in Drosophila. Curr Biol 7: 468-78.

Molyneaux, B. J., and G. M. Langford. 1997. Characterization of antibodies to the head and tail domains of squid brain myosin V. Biol Bull 193: 222-3.

Molyneaux, B. J., M. K. Mulcahey, P. Stafford, and G. M. Langford. 2000. Sequence and phylogenetic analysis of squid myosin-V: a vesicle motor in nerve cells. Cell Motil Cytoskeleton 46: 108-15.

Montell, C., and G. M. Rubin. 1988. The Drosophila ninaC locus encodes two photoreceptor cell specific proteins with domains homologous to protein kinases and the myosin heavy chain head. Cell 52: 757-72.

Montell, D. J., P. Rorth, and A. C. Spradling. 1992. slow border cells, a locus required for a developmentally regulated cell migration during oogenesis, encodes Drosophila C/EBP. Cell 71: 51-62.

Moore, K. J., J. R. Testa, U. Francke, A. Milatovich, N. G. Copeland, and N. A. Jenkins. 1995. Cloning and regional assignment of the human myosin heavy chain 12 (MYH12) gene to chromosome band 15q21. Cytogenet Cell Genet 69: 53-8.

Mooseker, M. S., and R. E. Cheney. 1995. Unconventional myosins. Annu Rev Cell Dev Biol 11: 633-75.

Morgan, N. S., D. M. Skovronsky, S. Artavanis-Tsakonas, and M. S. Mooseker. 1994. The molecular cloning and characterization of Drosophila melanogaster myosin-IA and myosin-IB. J Mol Biol 239: 347-56.

Morris, J., and R. Lehmann. 1999. Drosophila oogenesis: versatile spn doctors. Curr Biol 9: R55-8.

Muskavitch, M. A. 1994. Delta-notch signaling and Drosophila cell fate choice. Dev Biol 166: 415-30.

Nascimento, A. A., R. G. Amaral, J. C. Bizario, R. E. Larson, and E. M. Espreafico. 1997. Subcellular localization of myosin-V in the B16 melanoma cells, a wild- type cell line for the dilute gene. Mol Biol Cell 8: 1971-88.

Niki, I., H. Yokokura, T. Sudo, M. Kato, and H. Hidaka. 1996. Ca2+ signaling and intracellular Ca2+ binding proteins. J Biochem (Tokyo) 120: 685-98.

Norvell, A., R. L. Kelley, K. Wehr, and T. Schupbach. 1999. Specific isoforms of squid, a Drosophila hnRNP, perform distinct roles in Gurken localization during oogenesis. Genes Dev 13: 864-76.

Ohyama, A., Y. Komiya, and M. Igarashi. 2001. Globular tail of myosin-V is bound to vamp/synaptobrevin. Biochem Biophys Res Commun 280: 988-91.

O'Kane, C. J., and W. J. Gehring. 1987. Detection in situ of genomic regulatory elements in Drosophila. Proc Natl Acad Sci U S A 84: 9123-7.

Oliver, T. N., J. S. Berg, and R. E. Cheney. 1999. Tails of unconventional myosins. Cell Mol Life Sci 56: 243-57.

Onishi, H., S. Kojima, K. Katoh, K. Fujiwara, H. M. Martinez, and M. F. Morales. 1998. Functional transitions in myosin: formation of a critical salt-bridge and transmission of effect to the sensitive tryptophan. Proc Natl Acad Sci U S A 95: 6653-8.

Orr, W. C., V. K. Galanopoulos, C. P. Romano, and F. C. Kafatos. 1989. A female sterile screen of the Drosophila melanogaster X chromosome using hybrid dysgenesis: identification and characterization of egg morphology mutants. Genetics 122: 847-58.

Orr-Weaver, T. L. 1991. Drosophila chorion genes: cracking the eggshell's secrets. Bioessays 13: 97-105.

Orr-Weaver, T. L., and A. C. Spradling. 1986. Drosophila chorion gene amplification requires an upstream region regulating s18 transcription. Mol Cell Biol 6: 4624-33.

Osherov, N., and G. S. May. 2000. In vivo function of class I myosins. Cell Motil Cytoskeleton 47: 163-73.

Ostap, E. M., and T. D. Pollard. 1996. Overlapping functions of myosin-I isoforms? J Cell Biol 133: 221-4.

Page, R. D. M. 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci.: 357-358.

Park, I. C., H. Horiuchi, C. W. Hwang, W. H. Yeh, A. Ohta, J. C. Ryu, and M. Takagi. 1999. Isolation of csm1 encoding a class V chitin synthase with a myosin motor-like domain from the rice blast fungus, Pyricularia oryzae. FEMS Microbiol Lett 170: 131-9.

Patel, K. G., C. Liu, P. L. Cameron, and R. S. Cameron. 2001. Myr 8, a novel unconventional myosin expressed during brain development associates with the protein phosphatase catalytic subunits 1alpha and 1gamma1. J Neurosci 21: 7954-68.

Peifer, M., S. Orsulic, D. Sweeton, and E. Wieschaus. 1993. A role for the Drosophila segment polarity gene armadillo in cell adhesion and cytoskeletal integrity during oogenesis. Development 118: 1191-207.

Perrimon, N., and M. Gans. 1983. Clonal analysis of the tissue specificity of recessive female-sterile mutations of Drosophila melanogaster using a dominant female-sterile mutation Fs(1)K1237. Dev Biol 100: 365-73.

Porter, J. A., J. L. Hicks, D. S. Williams, and C. Montell. 1992. Differential localizations of and requirements for the two Drosophila ninaC kinase/myosins in photoreceptor cells. J Cell Biol 116: 683-93.

Prekeris, R., and D. M. Terrian. 1997. Brain myosin V is a synaptic vesicle-associated motor protein: evidence for a Ca2+-dependent interaction with the synaptobrevin-synaptophysin complex. J Cell Biol 137: 1589-601.

Provance, D. W., Jr., M. Wei, V. Ipe, and J. A. Mercer. 1996. Cultured melanocytes from dilute mutant mice exhibit dendritic morphology and altered melanosome distribution. Proc Natl Acad Sci U S A 93: 14554-8.

Reddy, A. S., and I. S. Day. 2001. Analysis of the myosins encoded in the recently completed Arabidopsis thaliana genome sequence. Genome Biol 2.

Reichelt, S., A. E. Knight, T. P. Hodge, F. Baluska, J. Samaj, D. Volkmann, and J. Kendrick-Jones. 1999. Characterization of the unconventional myosin VIII in plant cells and its localization at the post-cytokinetic cell wall. Plant J 19: 555-67.

Reinhard, J., A. A. Scheel, D. Diekmann, A. Hall, C. Ruppert, and M. Bahler. 1995. A novel type of myosin implicated in signalling by rho family GTPases. Embo J 14: 697-704.

Rhoads, A. R., and F. Friedberg. 1997. Sequence motifs for calmodulin recognition. Faseb J 11: 331-40.

Rief, M., R. S. Rock, A. D. Mehta, M. S. Mooseker, R. E. Cheney, and J. A. Spudich. 2000. Myosin-V stepping kinetics: a molecular model for processivity. Proc Natl Acad Sci U S A 97: 9482-6.

Robertson, H. M., C. R. Preston, R. W. Phillis, D. M. Johnson-Schlitz, W. K. Benz, and W. R. Engels. 1988. A stable genomic source of P element transposase in Drosophila melanogaster. Genetics 118: 461-70.

Robinson, D. N., and L. Cooley. 1997. Genetic analysis of the actin cytoskeleton in the Drosophila ovary. Annu Rev Cell Dev Biol 13: 147-70.

Robinson, D. N., K. Cant, and L. Cooley. 1994. Morphogenesis of Drosophila ovarian ring canals. Development 120: 2015-25.

Rodionov, V. I., A. J. Hope, T. M. Svitkina, and G. G. Borisy. 1998. Functional coordination of microtubule-based and actin-based motility in melanophores. Curr Biol 8: 165-8.

Rogers, S. L., R. L. Karcher, J. T. Roland, A. A. Minin, W. Steffen, and V. I. Gelfand. 1999. Regulation of melanosome movement in the cell cycle by reversible association with myosin V. J Cell Biol 146: 1265-76.

Rong, Y. S., S. W. Titen, H. B. Xie, M. M. Golic, M. Bastiani, P. Bandyopadhyay, B. M. Olivera, M. Brodsky, G. M. Rubin, and K. G. Golic. 2002. Targeted mutagenesis by homologous recombination in D. melanogaster. Genes Dev 16: 1568-81.

Rorth, P. 1998. Gal4 in the Drosophila female germline. Mech Dev 78: 113-8.

Roth, S., F. S. Neuman-Silberberg, G. Barcelo, and T. Schupbach. 1995. cornichon and the EGF receptor signaling process are necessary for both anterior-posterior and dorsal-ventral pattern formation in Drosophila. Cell 81: 967-78.

Santos, B., and M. Snyder. 1997. Targeting of chitin synthase 3 to polarized growth sites in yeast requires Chs5p and Myo2p. J Cell Biol 136: 95-110.

Schott, D., J. Ho, D. Pruyne, and A. Bretscher. 1999. The COOH-terminal domain of Myo2p, a yeast myosin V, has a direct role in secretory vesicle targeting. J Cell Biol 147: 791-808.

Schultz, J., F. Milpetz, P. Bork, and C. P. Ponting. 1998. SMART, a simple modular architecture research tool: identification of signaling domains. Proc Natl Acad Sci U S A 95: 5857-64.

Schupbach, T. 1987. Germ line and soma cooperate during oogenesis to establish the dorsoventral pattern of egg shell and embryo in Drosophila melanogaster. Cell 49: 699-707.

Schupbach, T., and E. Wieschaus. 1991. Female sterile mutations on the second chromosome of Drosophila melanogaster. II. Mutations blocking oogenesis or altering egg morphology. Genetics 129: 1119-36.

Schwarz, E. C., H. Geissler, and T. Soldati. 1999. A potentially exhaustive screening strategy reveals two novel divergent myosins in Dictyostelium. Cell Biochem Biophys 30: 413-35.

Segraves, W. A., and D. S. Hogness. 1990. The E75 ecdysone-inducible gene responsible for the 75B early puff in Drosophila encodes two new members of the steroid receptor superfamily. Genes Dev 4: 204-19.

Sellers, J. R. 2000. Myosins: a diverse superfamily. Biochim Biophys Acta 1496: 3-22.

Senda, S., S. F. Lee, G. P. Cote, and M. A. Titus. 2001. Recruitment of a specific amoeboid myosin I isoform to the plasma membrane in chemotactic Dictyostelium cells. J Biol Chem 276: 2898-904.

Sirotkin, V., S. Seipel, M. Krendel, and E. M. Bonder. 2000. Characterization of sea urchin unconventional myosins and analysis of their patterns of expression during early embryogenesis. Mol Reprod Dev 57: 111-26.

Solc, C. K., B. H. Derfler, G. M. Duyk, and D. P. Corey. 1994. Molecular cloning of myosins from bullfrog saccular macula: A candidate for the hair cell adaptation motor. Auditory Neurosci 1: 63-75.

Soldati, T., H. Geissler, and E. C. Schwarz. 1999. How many is enough? Exploring the myosin repertoire in the model eukaryote Dictyostelium discoideum. Cell Biochem Biophys 30: 389-411.

Soller, M., M. Bownes, and E. Kubli. 1999. Control of oocyte maturation in sexually mature Drosophila females. Dev Biol 208: 337-51.

Spradling, A. C. 1993. Developmental genetics of oogenesis, pp. 1-70. *In* M. Bate and A. M. Arias [eds.], The development of Drosophila melanogaster. Cold Spring Harbor Laboratory Press.

Spradling, A. C., and A. P. Mahowald. 1980. Amplification of genes for chorion proteins during oogenesis in Drosophila melanogaster. Proc Natl Acad Sci U S A 77: 1096-100.

Spradling, A. C., M. E. Digan, A. P. Mahowald, M. Scott, and E. A. Craig. 1980. Two clusters of genes for major chorion proteins of Drosophila melanogaster. Cell 19: 905-14.

Spradling, A. C., M. de Cuevas, D. Drummond-Barbosa, L. Keyes, M. Lilly, M. Pepling, and T. Xie. 1997. The Drosophila germarium: stem cells, germ line cysts, and oocytes. Cold Spring Harb Symp Quant Biol 62: 25-34.

Spradling, A. C., D. Stern, A. Beaton, E. J. Rhem, T. Laverty, N. Mozden, S. Misra, and G. M. Rubin. 1999. The Berkeley Drosophila Genome Project gene disruption project: Single P-element insertions mutating 25% of vital Drosophila genes. Genetics 153: 135-77.

St Johnston, D., D. Beuchle, and C. Nusslein-Volhard. 1991. Staufen, a gene required to localize maternal RNAs in the Drosophila egg. Cell 66: 51-63.

St Johnston, D., W. Driever, T. Berleth, S. Richstein, and C. Nusslein-Volhard. 1989. Multiple steps in the localization of bicoid RNA to the anterior pole of the Drosophila oocyte. Development 107: 13-9.

Stephenson, E. C., Y. C. Chao, and J. D. Fackenthal. 1988. Molecular analysis of the swallow gene of Drosophila melanogaster. Genes Dev 2: 1655-65.

Stevens, R. C., and T. N. Davis. 1998. Mlc1p is a light chain for the unconventional myosin Myo2p in Saccharomyces cerevisiae. J Cell Biol 142: 711-22.

Suter, D. M., F. S. Espindola, C. H. Lin, P. Forscher, and M. S. Mooseker. 2000. Localization of unconventional myosins V and VI in neuronal growth cones. J Neurobiol 42: 370-82.

Sutherland, J. D., T. Kozlova, G. Tzertzinis, and F. C. Kafatos. 1995. Drosophila hormone receptor 38: a second partner for Drosophila USP suggests an unexpected role for nuclear receptors of the nerve growth factor-induced protein B type. Proc Natl Acad Sci U S A 92: 7966-70.

Swan, A., and B. Suter. 1996. Role of Bicaudal-D in patterning the Drosophila egg chamber in mid- oogenesis. Development 122: 3577-86.

Swan, A., S. Hijal, A. Hilfiker, and B. Suter. 2001. Identification of new X-chromosomal genes required for Drosophila oogenesis and novel roles for fs(1)Yb, brainiac and dunce. Genome Res 11: 67-77.

Swanson, J. A., M. T. Johnson, K. Beningo, P. Post, M. Mooseker, and N. Araki. 1999. A contractile activity that closes phagosomes in macrophages. J Cell Sci 112: 307-16.

Tautz, D., J. M. Hancock, D. A. Webb, C. Tautz, and G. A. Dover. 1988. Complete sequences of the rRNA genes of Drosophila melanogaster. Mol Biol Evol 5: 366-76.

Theodosiou, N. A., and T. Xu. 1998. Use of FLP/FRT system to study Drosophila development. Methods 14: 355-65.

Theurkauf, W. E. 1994. Microtubules and cytoplasm organization during Drosophila oogenesis. Dev Biol 165: 352-60.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-80.

Thummel, C. S., and V. Pirrotta. 1992. New pCaSpeR P element vectors. Dros. Info Services 71: 150.

Titus, M. A. 1997. Motor proteins: myosin V--the multi-purpose transport motor. Curr Biol 7: R301-4.

Titus, M. A. 1999. A class VII unconventional myosin is required for phagocytosis. Curr Biol 9: 1297-303.

Torok, T., G. Tick, M. Alvarado, and I. Kiss. 1993. P-lacW insertional mutagenesis on the second chromosome of Drosophila melanogaster: isolation of lethals with different overgrowth phenotypes. Genetics 135: 71-80.

Tzolovsky, G., W. M. Deng, T. Schlitt, and M. Bownes. 1999. The function of the broad-complex during Drosophila melanogaster oogenesis. Genetics 153: 1371-83.

Tzolovsky, G., H. Millo, S. Pathirana, T. Wood, and M. Bownes. 2002. Identification and Phylogenetic Analysis of Drosophila melanogaster Myosins. Mol Biol Evol 19: 1041-52.

van Eeden, F., and D. St Johnston. 1999. The polarisation of the anterior-posterior and dorsal-ventral axes during Drosophila oogenesis. Curr Opin Genet Dev 9: 396-404.

Vargas, M. A., H. Voigt, P. Sansonetti, and N. Guillen. 1997. The tail domain of Entamoeba histolytica myosin IB bind F-actin. Arch Med Res 28: 137-8.

Volkmann, N., and D. Hanein. 2000. Actomyosin: law and order in motility. Curr Opin Cell Biol 12: 26-34.

von Kalm, L., K. Crossgrove, D. Von Seggern, G. M. Guild, and S. K. Beckendorf. 1994. The Broad-Complex directly controls a tissue-specific response to the steroid hormone ecdysone at the onset of Drosophila metamorphosis. Embo J 13: 3505-16.

Wang, F., L. Chen, O. Arcucci, E. V. Harvey, B. Bowers, Y. Xu, J. A. Hammer, 3rd, and J. R. Sellers. 2000. Effect of ADP and ionic strength on the kinetic and motile properties of recombinant mouse myosin V. J Biol Chem 275: 4329-35.

Wang, F. S., J. S. Wolenski, R. E. Cheney, M. S. Mooseker, and D. G. Jay. 1996. Function of myosin-V in filopodial extension of neuronal growth cones. Science 273: 660-3.

Wang, S., and T. Hazelrigg. 1994. Implications for bcd mRNA localization from spatial distribution of exu protein in Drosophila oogenesis. Nature 369: 400-03.

Waring, G. L., and A. P. Mahowald. 1979. Identification and time of synthesis of chorion proteins in Drosophila melanogaster. Cell 16: 599-607.

Wes, P. D., X. Z. Xu, H. S. Li, F. Chien, S. K. Doberstein, and C. Montell. 1999. Termination of phototransduction requires binding of the NINAC myosin III and the PDZ protein INAD. Nat Neurosci 2: 447-53.

Wirth, J. A., K. A. Jensen, P. L. Post, W. M. Bement, and M. S. Mooseker. 1996. Human myosin-IXb, an unconventional myosin with a chimerin-like rho/rac GTPase-activating protein domain in its tail. J Cell Sci 109: 653-61.

Wolenski, J. S. 1995. Regulation of calmodulin-binding myosins. Trends Cell Biology 5: 310-316.

Wolenski, J. S., R. E. Cheney, M. S. Mooseker, and P. Forscher. 1995. In vitro motility of immunoadsorbed brain myosin-V using a Limulus acrosomal process and optical tweezer-based assay. J Cell Sci 108: 1489-96.

Wolfe, K. H., and D. C. Shields. 1997. Molecular evidence for an ancient duplication of the entire yeast genome. Nature 387: 708-13.

Wu, X., G. Jung, and J. A. Hammer, 3rd. 2000. Functions of unconventional myosins. Curr Opin Cell Biol 12: 42-51.

Wu, X., B. Kocher, Q. Wei, and J. A. Hammer, 3rd. 1998. Myosin Va associates with microtubule-rich domains in both interphase and dividing cells. Cell Motil Cytoskeleton 40: 286-303.

Wu, X., B. Bowers, Q. Wei, B. Kocher, and J. A. Hammer, 3rd. 1997. Myosin V associates with melanosomes in mouse melanocytes: evidence that myosin V is an organelle motor. J Cell Sci 110: 847-59.

Xu, P., K. I. Mitchelhill, B. Kobe, B. E. Kemp, and H. G. Zot. 1997. The myosin-I-binding protein Acan125 binds the SH3 domain and belongs to the superfamily of leucine-rich repeat proteins. Proc Natl Acad Sci U S A 94: 3685-90.

Yamashita, R. A., J. R. Sellers, and J. B. Anderson. 2000. Identification and analysis of the myosin superfamily in Drosophila: a database approach. J Muscle Res Cell Motil 21: 491-505.

Yao, T. P., W. A. Segraves, A. E. Oro, M. McKeown, and R. M. Evans. 1992. Drosophila ultraspiracle modulates ecdysone receptor function via heterodimer formation. Cell 71: 63-72.

Zaccai, M., and H. D. Lipshitz. 1996. Role of Adducin-like (hu-li tai shao) mRNA and protein localization in regulating cytoskeletal structure and function during Drosophila Oogenesis and early embryogenesis. Dev Genet 19: 249-57.

Zelhof, A. C., T. P. Yao, R. M. Evans, and M. McKeown. 1995. Identification and characterization of a Drosophila nuclear receptor with the ability to inhibit the ecdysone response. Proc Natl Acad Sci U S A 92: 10477-81.

Zelhof, A. C., N. Ghbeish, C. Tsai, R. M. Evans, and M. McKeown. 1997. A role for ultraspiracle, the Drosophila RXR, in morphogenetic furrow movement and photoreceptor cluster formation. Development 124: 2499-506.

Zhang, S., and S. I. Bernstein. 2001. Spatially and temporally regulated expression of myosin heavy chain alternative exons during Drosophila embryogenesis. Mech Dev 101: 35-45.

Zhao, L. P., J. S. Koslovsky, J. Reinhard, M. Bahler, A. E. Witt, D. W. Provance, Jr., and J. A. Mercer. 1996. Cloning and characterization of myr 6, an unconventional myosin of the dilute/myosin-V family. Proc Natl Acad Sci U S A 93: 10826-31.

Zhimulev, I. F., E. S. Belyaeva, O. M. Mazina, and M. L. Balasov. 1995. Structure and expression of the br-c locus in Drosophila melanogaster (Diptera: Drosophilidae). Eur J Entomol 92: 263-270.

Zollman, S., D. Godt, G. G. Prive, J. L. Couderc, and F. A. Laski. 1994. The BTB domain, found primarily in zinc finger proteins, defines an evolutionarily conserved family that includes several developmentally regulated genes in Drosophila. Proc Natl Acad Sci U S A 91: 10717-21.

The Function of the *Broad-Complex* During *Drosophila melanogaster* Oogenesis

George Tzolovsky,[1] Wu-Min Deng,[1,2] Thomas Schlitt[3] and Mary Bownes

Institute of Cell and Molecular Biology, University of Edinburgh, Edinburgh EH9 3JR, United Kingdom

Manuscript received May 7, 1999
Accepted for publication July 30, 1999

ABSTRACT

The *Broad-Complex* (*BR-C*) is an early ecdysone response gene that functions during metamorphosis and encodes a family of zinc-finger transcription factors. It is expressed in a dynamic pattern during oogenesis. Its late expression in the lateral-dorsal-anterior follicle cells is related to the morphogenesis of the chorionic appendages. All four zinc-finger isoforms are expressed in oogenesis, which is consistent with the abnormal appendage phenotypes resulting from their ectopic expression. We investigated the mechanism by which the *BR-C* affects chorion deposition by using BrdU to follow the effects of *BR-C* misexpression on DNA replication and *in situ* hybridization to ovarian mRNA to evaluate chorion gene expression. Ectopic *BR-C* expression leads to prolonged endoreplication and to additional amplification of genes, besides the chorion genes, at other sites in the genome. The pattern of chorion gene expression is not affected along the anterior-posterior axis, but the follicle cells at the anterior of the oocyte fail to migrate correctly in an anterior direction when *BR-C* is misexpressed. We conclude that the target genes of the *BR-C* in oogenesis include a protein essential for endoreplication and chorion gene amplification. This may provide a link between steroid hormones and the control of DNA replication during oogenesis.

THE regulation of the expression of structural genes is critical in morphogenesis. This requires differential expression of transcription factors, which in turn regulate the tissue-specific expression of structural genes. Drosophila oogenesis is ideal for the study of developmental gene regulation as it takes a fairly short time to develop from a stage-1 egg chamber to a mature egg and all the stages are morphologically well defined. Further, egg chambers, the developmental units of oogenesis, contain only the somatically derived follicle cells and the germline cells. The former undergo dramatic morphogenetic movements and eventually synthesize the yolk and then the eggshell, as well as interact with the germline cells to generate the two major axes of the egg and embryo.

The *Broad-Complex* (*BR-C*), a gene encoding a family of zinc-finger transcription factors (DiBello *et al.* 1991; Bayer *et al.* 1996), has been shown to be expressed in the follicle cells in a dynamic pattern, the late pattern being defined by two groups of dorsal-anterior follicle cells at stage 10B of oogenesis (Deng and Bownes 1997; for the staging of oogenesis, refer to Spradling 1993). This dorsal-anterior expression pattern is specified by the Grk-DER and decapentaplegic (DPP) signaling pathways along the two major axes, and is associated with the function of the *BR-C* in dorsal appendage formation. The involvement of *BR-C* in dorsal appendage morphogenesis was shown by mutational analysis of *BR-C* partial "loss-of-function" mutants, and was supported by ectopic expression of *BR-C* "transgenes" during oogenesis. It is proposed that the *BR-C* may provide a link between pattern formation and cell behavior in morphogenesis (Deng and Bownes 1997).

The *BR-C* has been previously identified as a key gene required for Drosophila metamorphosis. It is among the early ecdysone responsive genes, which are directly activated by the ecdysone receptor and coordinate the subsequent transcription of the tissue-specific "late genes" (Ashburner 1974; and for reviews see Kiss *et al.* 1988; Karim *et al.* 1993; Zhimulev *et al.* 1995; Bayer *et al.* 1996a). The *BR-C* is located at chromosomal region 2B5. Genetically, the *BR-C* locus has three fully complementing functions: *br* (*broad*), *rbp* (reduced bristle number on palpus), and *2Bc*, as well as a noncomplementing *npr* (nonpupariating) class (Figure 1C; Belyaeva *et al.* 1980). Additionally, a number of *BR-C* alleles have been categorized to the *2Bab* group. These alleles do not complement *br* or *rbp* mutations, but do complement *2Bc* mutations (Belyaeva *et al.* 1980). The nonpupariating mutations are probably null mutations, because alleles in this class fail to complement mutations in each of the three complementing groups. These mutants are also phenotypically indistinguishable from deletions of the locus. It has been shown by genetic analysis that *BR-C* is essential for the morphogenesis of imaginal discs. br^+ function is primarily required in the elonga-

Corresponding author: Mary Bownes, Institute of Cell and Molecular Biology, University of Edinburgh, Darwin Bldg., King's Bldgs., Mayfield Rd., Edinburgh EH9 3JR, United Kingdom.
E-mail: mary.bownes@ed.ac.uk

[1] These authors contributed equally to this work.

[2] *Present address:* Department of Biochemistry, Box 357350, University of Washington, Seattle, WA 98195.

[3] *Present address:* Genetisches Institut der Justus-Liebig-Universitaet, Heinrich-Buff-Ring 58-62, D-35392, Giessen, Germany.

tion and eversion of appendages from imaginal discs as well as tanning and hardening of the larval cuticle. rbp^+ function, on the other hand, is essential for muscle and bristle development. Additionally, both rbp^+ and $2Bc^+$ functions are vital for histolysis of the larval tissues and gut morphogenesis. $2Bc^+$ function was shown to be essential in the fusion of discs to form a continuous adult epidermis (Kiss *et al.* 1988). All three functions are also required for the reorganization of the central nervous system (CNS) (Kiss *et al.* 1988; Emery *et al.* 1994).

The *BR-C* encodes a family of C_2H_2 zinc-finger proteins (Z1, Z2, Z3, and Z4), which share a common aminoterminal (the *BR-C* "core") domain but differ in zinc-finger domains (DiBello *et al.* 1991; Bayer *et al.* 1996b). The core contains a highly conserved amino-terminal motif, called the BTB or POZ domain, which appears to be involved in protein-protein interactions and is widely distributed among metazoans (DiBello *et al.* 1991; Bardwell and Treisman 1994; Zollman *et al.* 1994). The core is alternatively spliced to one of the four zinc-finger domains (Figure 1), generating four classes of proteins, the Z1, Z2, Z3, and Z4 isoforms. Additionally, three variants of the Z1 isoform have been identified. They differ in the linker region between the core motif and the zinc-finger domain.

Some genetic studies suggest a one-to-one link be-

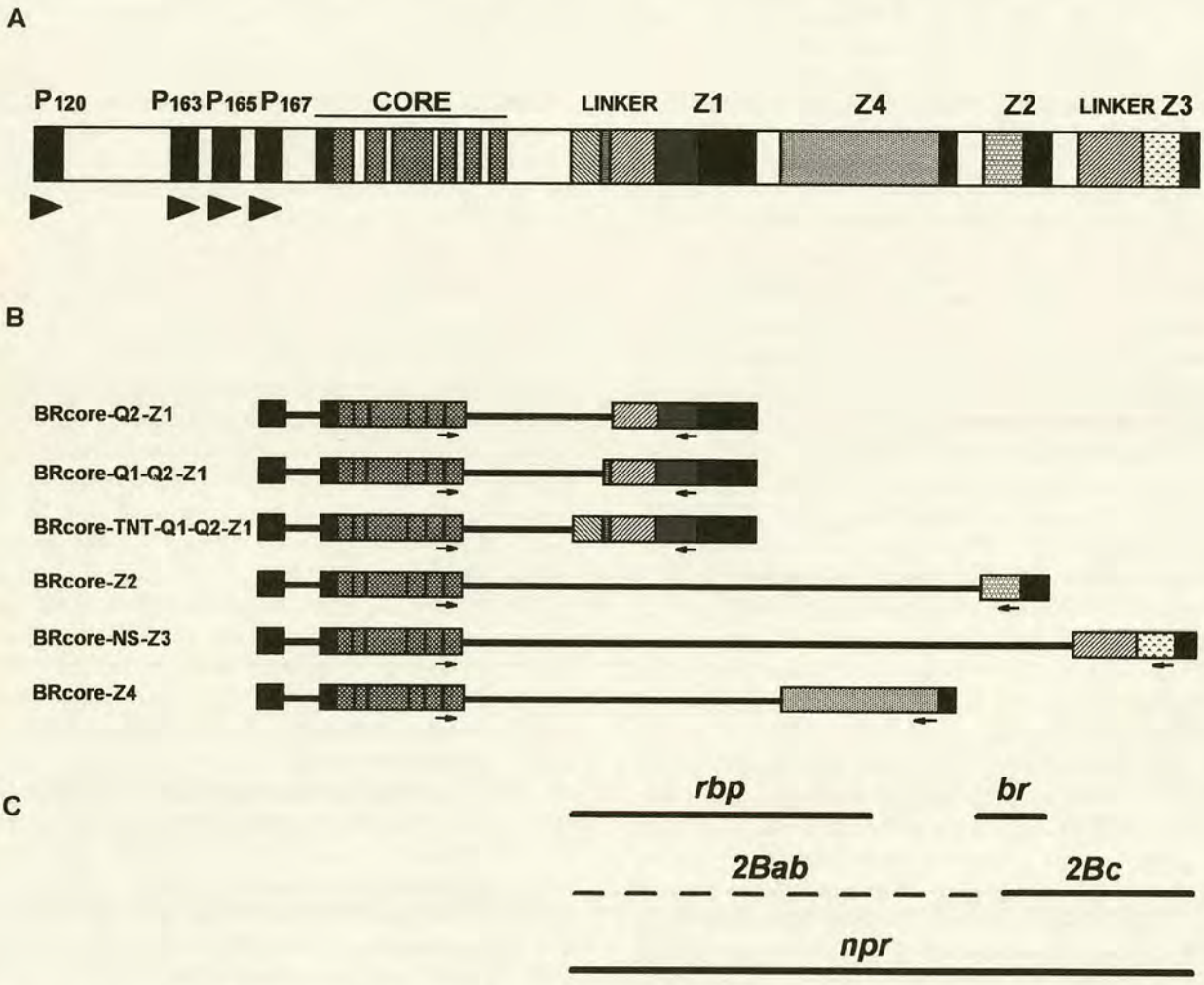


Figure 1.—Organization of the *BR-C* gene. (A) Molecular organization of the *BR-C* gene, which maps to the 2B5 region (adapted from Chao and Guild 1986). For simplicity and clarity the upstream half of *BR-C*, including the first exon, is not shown in the figure. Shaded boxes represent open reading frames; solid boxes represent untranslated regions of the *BR-C* transcripts. Two putative promoters have been previously described: P distal at nucleotide 120; and P proximal at nucleotides 163, 165, and 167 (DiBello *et al.* 1991; Bayer *et al.* 1996b). The Z1–Z2 transcripts can be synthesized from either P_{120} or P_{165}, while Z2 and Z4 initiate only at P_{165} and P_{163}, respectively (Bayer *et al.* 1996b). The linker regions for Z1 and Z3 domains are contiguous in the respective exon. (B) Organization of zinc-finger isoforms. Differentially spliced *BR-C* transcripts share a common core domain linked to one of the four (Z1–Z4) different pairs of C2H2 zinc-finger domains. The TNT, Q_1, and Q_2 linker sequences found in Z1 transcripts are contiguous in the Z1 exon, and generate by alternatively splicing three Z1 isoforms (Bayer *et al.* 1997). Arrows indicate the primer pairs used in RT-PCR analysis. (C) Complementation map of the *BR-C* based on Bayer *et al.* (1997).

tween the specific complementing genetic functions and protein isoforms. However, other data suggest that the relationships between the complementing groups and protein isoforms are more complicated. For example, in the br^{28} mutant, Z3 transcripts and protein levels are reduced and all Z1 isoforms are truncated. Clearer data on these relationships were provided by BAYER *et al.* (1997), who showed that lethality associated with each of the complementing groups was rescued using heat-inducible transgenes. It was found that br^+ function is only provided by the Z2 isoform. Despite this, there may be functional redundancy or regulatory dependency associated with rbp^+ and $2Bc^+$ functions. It was found that Z1 transgenes provide full rbp^+ function, while Z4 provides partial function. The *2Bc* lethality is fully rescued by Z3 protein expression, and partially rescued by Z2 protein expression.

The two clusters of chorion genes on the X-chromosome and third chromosome, which are responsible for the production of large amounts of chorion protein in the follicle cells at very precisely defined points in late oogenesis, are selectively amplified above the level of the remainder of the follicle cell genome, which also endoreplicates to produce polyploid cells (SPRADLING and MAHOWALD 1980; ORR-WEAVER 1991). The third chromosome cluster is amplified 60- to 80-fold and the X-chromosome group 15- to 20-fold above the rest of the genome. This specific amplification depends on *cis*-acting sequences among the chorion genes (ORR-WEAVER and SPRADLING 1986; DELIDAKIS and KAFATOS 1989).

We know that the *BR-C* expression in the lateral-dorsal-anterior follicle cells during oogenesis is related to its function in dorsal appendage formation. However, we do not know what the early function of the *BR-C* is, when it is expressed in all follicle cells at stage 6 of oogenesis. Since the chorion genes encode major eggshell components, and rbp^+ function has been reported to be necessary for chorion gene amplification during oogenesis (ORR *et al.* 1989; HUANG and ORR 1992), we investigate in this article the relationship between the *BR-C* and chorion gene amplification and expression.

MATERIALS AND METHODS

Drosophila strains: The following *Drosophila melanogaster* strains were used: Oregon R, br^1, br^5, br^6, rbp^1, rbp^2, $2Bc^1$, $2Bc^2$, npr^6 (KISS *et al.* 1988); br^{A47} (DENG and BOWNES 1997); w^{1118}, *hs.Z1* (527-5; 708-1), *hs.Z2* (CD5-1), *hs.Z3* (797-3), *hs.Z4* (Z4-11) (BAYER *et al.* 1997). The rbp^2 and Br^1 alleles are viable and were maintained as homozygotes. All other *BR-C* mutations were maintained over *Binsn*, an X-chromosome balancer carrying the markers *Bar* and *singed*. All stocks were maintained on standard cornmeal food at 26°.

Antibody staining of ovaries: Ovaries were dissected from yeasted flies in Ringer's solution. The anterior parts of the ovaries were torn apart to facilitate antibody penetration. The ovaries were transferred to a microfuge tube containing 2% *p*-formaldehyde (in 1 × PBS) and fixed for 30 min at room temperature. The fixative was carefully removed and ovaries were washed in 1 ml of PTW [1.5% (v/v) Tween-20 in PBS] for 5 min. Then the ovaries were incubated in 1% (w/v) bovine serum albumin (Sigma, St. Louis) in PTW for 1 hr. Blocking was accomplished by incubation of the ovaries in PTW-NGS [5% (v/v) normal goat serum in PTW] for 2 hr at room temperature. The first antibody was added at 1:200 dilution in PTW and the incubation was carried out overnight at 4°. Residual antibody was washed away with three changes of PTW with 30 min of incubation per change. The HRP-conjugated secondary antibody (Promega, Madison, WI) was then added to the ovaries at 1:500 dilution and incubated for 2 hr at room temperature or overnight at 4°. Excess secondary antibody was removed with three PTW washes at 30-min intervals. Diaminobenzidine (DAB) staining solution (Sigma) was added and the staining was allowed to proceed for 10–30 min before washing with several changes of PBS to stop the reaction. Stained ovaries were mounted in PBS/glycerol (1:4) to allow microscopy.

Hoechst staining: Ovaries were dissected in PBS and fixed in 4% *p*-formaldehyde (w/v in 1 × PBS) for 20 min. This was followed by washing in 1 × PBT [1% (v/v) Triton-X100 in PBS] for 30 min. The ovaries were then washed for 30 min in PBS and stained for 5 min in 1 μg/ml Hoechst 33258 (Sigma; dissolved in PBS). After washing in PBS for 2 hr to overnight, the ovaries were mounted in PBS/glycerol (1:4) and examined under a fluorescent microscope.

Preparation of the eggshell for dark-field microscopy: Freshly laid eggs were collected from the apple juice plate and placed in a drop of Hoyer's mountant (Hoyer's mounting medium:lactic acid = 1:1) on a glass slide and covered by a coverslip. After an overnight incubation at 65° the slides were ready for dark-field microscopy.

RNA extraction and RT-PCR: The *BR-C* transcript levels in ovaries were detected by reverse transcriptase (RT)-PCR as described previously (HODGETTS *et al.* 1995). Total RNA from ovaries and larvae (control RNA) was isolated using RNAeasy-Total RNA Kit (QIAGEN, Chatsworth, CA, no. 74104). The RNA (5 μg) was primed with oligo-p$(dT)_{15}$ and reverse transcribed using Superscript II (Gibco BRL, Gaithersburg, MD) following the supplier's protocol. For the subsequent DNA amplification, 5% of the first-strand reaction mix was used. To amplify each of the zinc-finger domains, appropriate primer pairs were added to the PCR mixture: a common primer for the core domain was combined with one of the four primers for the respective zinc-finger motif. The sequence data for the primers were obtained from DIBELLO *et al.* (1991) and HODGETTS *et al.* (1995): core, 5′-ACAAGATGTTCCATG CAGCC-3′; Z1, 5′-TGCTGGTGCTGCTGGTGATA-3′; Z2, 5′-TCATCTCCATTTCGCCGGGA-3′; Z3, 5′-GATGGCGGTCGT CTTAAGCA-3′; Z4, 5′-GTGGTTGTTCAGCGAGTTCA-3′. In the PCR reaction QIAGEN Taq Polymerase and the protocol designed for use with Q-Solution was used. The PCR reaction was carried out as follows: one cycle at 94° for 4 min; 35 cycles, step one at 94° for 30 sec, step two at 60° for 30 sec, step three at 72° for 1.5 min; one cycle at 72° for 7 min.

BrdU labeling: Ovaries were dissected at room temperature in 1 × Grace's medium (Flow Laboratories, no. 2700049) and incubated for 1 hr in 15 μM BrdU (Sigma) in Grace's medium (LILLY and SPRADLING 1996). After washing in EBR (Ephrussi Beadle Ringer) the ovaries were fixed for 20 min in 37% formaldehyde/buffer B/dH_2O (1:1:4; LIN and SPRADLING 1993), followed by 1 hr denaturing in 2 N HCl and neutralizing for 2 min in 100 mM Na tetraborate. The tissue was rinsed several times in PBT (PBS + 0.1% Triton X-100) and blocked for 1 hr in 5% NGS in PBT. After overnight incubation in 1:20 dilution of anti-BrdU antibody (Becton-Dickinson, San Jose, CA, no. 347580) detection was carried out with HRP-

conjugated secondary antibody (1:25 dilution). Sigma Fast DAB peroxidase substrate (no. 4168) was used in the peroxidase color reaction. The latter was enhanced with 10 μl 1 M Ni SO_4 per 1 ml staining solution.

***In situ* hybridization to mRNA in ovaries:** The protocol is based on the procedure previously described (Tautz and Pfeifle 1998) and modified as follows. The ovaries were dissected in Ringer's solution and fixed for 20 min in 4% *p*-formaldehyde in PBS. After rinsing the tissue in PBT it was treated for 10 min in methanol/0.5 M EGTA, pH 8 (9:1). The ovaries can then be stored in methanol at −20° for several months. The stored ovaries were rehydrated in PBT. The prehybridization was carried out for 1 hr at 45° in DNA Hybrix (50% deionized formamide, 5× SSC, 100 μg/ml sonicated salmon sperm DNA, 50 μg/ml Heparin, 0.1% Tween 20). The ovaries were hybridized overnight at 45° in DNA Hybrix containing digoxigenin-labeled probe (DIG DNA labeling and detection kit, Boehringer Mannheim, Indianapolis). For detection a 1:1000 dilution of anti-DIG-AP-conjugated Ab was used. The staining reaction was performed in 100 mM Tris pH 9.5, 50 mM MgCl, 10 mM NaCl, 0.2% Tween 20, 8 mM levamisole, 4.5 μl/ml NBT, and 3.5 μl/ml X-phosphate (Boehringer Mannheim). Anti-DIG-AP conjugate was preabsorbed with postfixed wild-type (Oregon R) ovaries at 4° overnight. The ovaries were mounted in a mixture of PBS/glycerol (1:4) for microscopy.

RESULTS

The *BR-C* protein distribution pattern during oogenesis: In a previous article (Deng and Bownes 1997), we reported that *BR-C* mRNA is expressed in follicle cells in a dynamic pattern. Its expression is first detected in all follicle cells at stage 6. During stage 10A, all the columnar cells, except the dorsal anterior follicle cells, contain the *BR-C* transcript. However, only two groups of dorsal-lateral-anterior follicle cells express the *BR-C* mRNA during stage 10B, marking the dorsal appendage secreting cells.

To investigate the function of the *BR-C* we need to establish whether or not the protein is distributed in a similar pattern to the mRNA during oogenesis. Antibodies that recognize the BR-core, Z1, or Z3 domains, respectively, were used to stain the whole-mount ovaries. Antibodies to the Z2 and Z4 isoforms have not been generated, so we were unable to check their expression pattern. Both the Z1 and BR-core antibodies exhibited similar staining patterns, while the Z3 antibody showed no staining during oogenesis. These observations are consistent with the results shown by RNA *in situ* hybridization; Z1 is the only zinc-finger isoform with expression at levels significantly high to be detected by *in situ* hybridization techniques during oogenesis.

The distribution pattern of the *BR-C* protein appears to be similar to that of its mRNA during stages 6–8 of oogenesis, when all follicle cells stain (Figure 2, A1). The protein is also detected in all columnar follicle cells except the dorsal anterior cells at stage 10 (Figure 2, A1–A3), similar to the pattern of mRNA distribution (Deng and Bownes 1997). However, the follicle cells at the posterior pole appear to be stained at this stage (Figure 2, A1), which differs from the mRNA distribution pattern. The late distribution pattern of the *BR-C* protein and mRNA differs. A very strong signal is observed in two groups of the lateral-dorsal follicle cells at stages 11 and 12, but the posterior and ventral follicle cells are still stained (Figure 2B). The signal in the posterior and ventral region disappears around late stage 13, leaving only the dorsal-appendage-associated

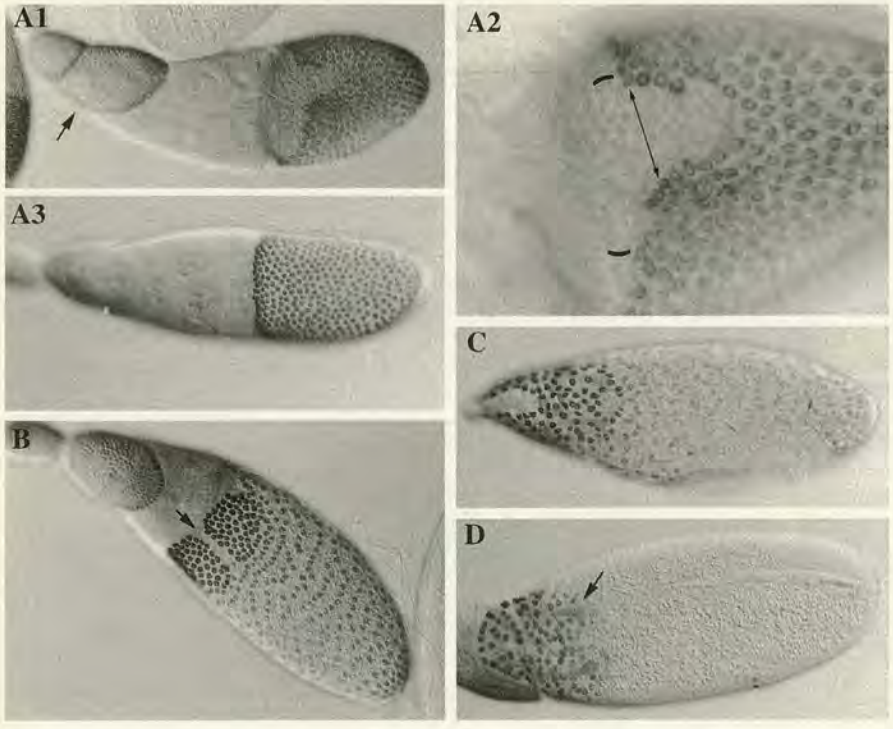


Figure 2.—The distribution pattern of the *BR-C* protein during oogenesis. (A1) Using BR-core antibody to stain the ovaries, signals are initially detected in all follicle cells at around stage 6 (arrowhead). Staining is also observed in all follicle cells at stage 7. During stage 10, columnar follicle cells over the oocyte are stained. However, staining is not seen at the dorsal anterior region. (A2) A closer look at the dorsal region. The dorsal gap is marked by a double-headed arrow, while the anterior gap is labeled by two curved lines. (A3) A ventral view of the same egg chamber. It appears that all ventral columnar cells are stained. (B) During stage 12, two groups of lateral-dorsal-anterior follicle cells are heavily stained, while the posterior and the ventral follicle cells are weakly stained. The dorsal gap between the two groups of the lateral-dorsal cells still exists (arrow). Expression in the posterior follicle cells becomes gradually weaker (C), until it disappears at approximately stage 13 (D). Expression is only detectable in the dorsal-appendage-associated follicle cells. The arrow shows a growing dorsal appendage.

follicle cells stained (Figure 2, C and D). The differences between the distribution patterns of the protein and mRNA presumably reflect the fact that the half-life of the protein is much longer than that of the mRNA. By the time late *BR-C* transcription occurs at the lateral-dorsal-anterior follicle cells, the protein translated from the early *BR-C* transcripts remains at the posterior and ventral side, while the mRNA has been degraded. Thus, the early and late protein distribution patterns overlap to form a gradient-like pattern at stages 11 and 12. The same reasoning could also be used to explain why the protein, but not the mRNA, is detected in the follicle cells at the posterior pole during stage 10.

Another feature of the *BR-C* protein distribution is that it only appears in the nuclei of the follicle cells (Figure 2), which is consistent with the fact that the *BR-C* encodes transcription factors.

rbp^+ function is required for dorsal appendage formation: The genetic organization of the *BR-C* is shown in Figure 1. Bayer *et al.* (1997) reported that Z1 provides the full rbp^+ function. Since the Z1 zinc-finger isoform is expressed during oogenesis and is detectable by *in situ* hybridization, it is predicted that the *rbp* functional domain will be required. To test this, female homozygous viable rbp^1 and rbp^2 mutants were dissected to examine the eggshell phenotype. It was found that the dorsal appendages were abnormal, being shorter and rougher than the wild type, and the eggshells were much more fragile (Figure 3). Shortening of the dorsal appendages was observed in rbp^2 homozygous mutants.

To test if *rbp* is the major functional domain involved in dorsal appendage formation we did a genetic analysis using the *npr* allele in crosses with *br*, *rbp*, and *2Bc* alleles (Table 1). The cross between rbp^1 and npr^6 produced only two males and no viable female heteroallelic mutants. Two rbp^2/npr^6 females were produced in the cross with rbp^2 flies. They lived for 2 days without laying any eggs. Then they were dissected to examine the ovarian phenotype. It was found that the ovaries were not completely developed and the few late stage oocytes formed had no appendages. Thus we were unable to examine *rbp* function by this method. The combination $2Bc/npr^6$ was found to be completely lethal, and as a result we could not establish if there is a function encoded by the Z3 (2Bc) isoform.

The function of Z2 was investigated by further genetic crosses. The cross between br^5 and npr^6 produced no viable heteroallelic flies, while the cross between br^6 and npr^6 generated 22 heteroallelic males, but no females that could be examined. It was observed that eggs produced by br^1/br^{A47} and br^1/npr^6 mothers have reduced dorsal appendages (Deng and Bownes 1997), suggesting that the *br* functional domain is likely to be required for dorsal appendage formation. To test this hypothesis, eggs produced by br^5/br^1 and br^6/br^1 females were examined, and found to have normal dorsal appendages (data not shown). This observation, along with the fact that the br^{del2} is actually an *rbp* allele (Huang and Orr 1992), suggests that the *br* functional domain, and hence Z2, is not involved in dorsal appendage formation.

How can the phenotype of br^1/br^{A47} and br^1/npr^6 eggshells be explained if *br* is not the functional domain required for dorsal appendage formation? This could be understood if the br^1 mutant not only affected *br* function, but also affected *rbp* function. To test this possibility, the eggshell phenotype of eggs laid from br^1/rbp^1 mutants was examined. It was shown that eggs produced by the br^1/rbp^1 mothers have reduced dorsal appendages, similar to those produced by the br^1/br^{A47} females. This indicates, therefore, that the br^1 is in fact a weak *2Bc* or *2Bab* allele, which fails to complement either *rbp* or *br* function. This suggests that *rbp* (which encodes Z1 and Z4) is a functional domain involved in dorsal appendage formation during oogenesis; however, we cannot rule out the involvement of Z3 from these experiments due to the failure of these crosses to generate adult females due to early lethality.

Ectopic *BR-C* expression induces ectopic dorsal appendage material: Is *BR-C* function sufficient to direct

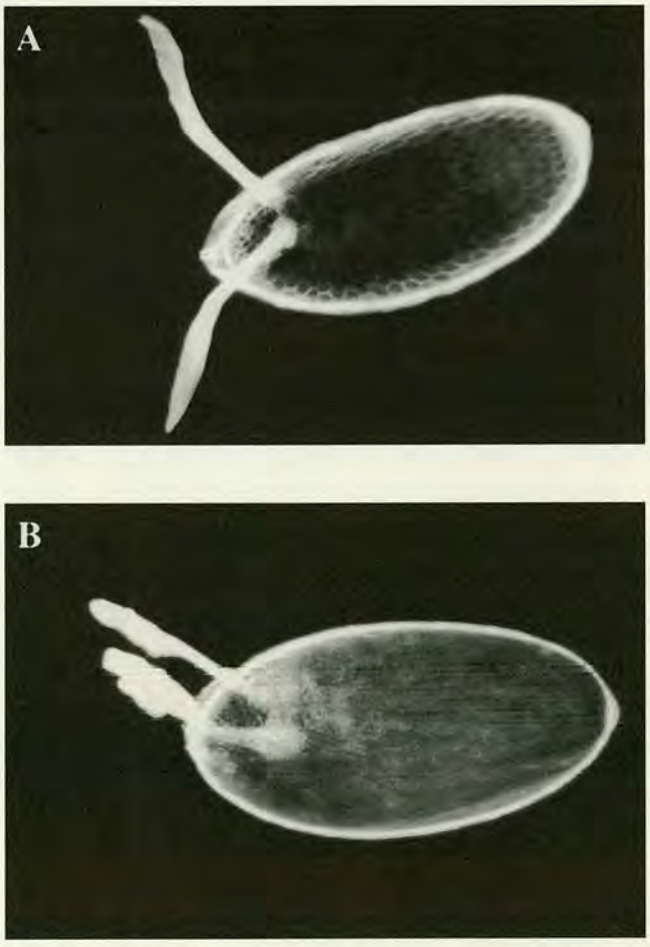


Figure 3.—Chorion defects in *rbp* mutants. (A) Wild-type chorionic appendages. (B) Appendages that are shorter and irregular in shape, produced by rbp^1 females.

TABLE 1

Survival frequencies of *BR-C* mutants

Cross	Progeny: Heterozygotes[a]: +Homozygous balancer[b]	Heteroallelic mutants[c] ♂	Heteroallelic mutants[c] ♀	Lethal period
$\frac{br^1}{br^1} \times \frac{npr^6}{Binsn}$	216	21	9	Not observed
$\frac{br^5}{Binsn} \times \frac{npr^6}{Binsn}$	158	0	0	III instar larva
$\frac{br^6}{Binsn} \times \frac{npr^6}{Binsn}$	36	22	0	Pupal
$\frac{rbp^1}{Binsn} \times \frac{npr^6}{Binsn}$	140	2	0	Pupal
$\frac{rbp^2}{rbp^2} \times \frac{npr^6}{Binsn}$	130	11	2	Pupal
$\frac{2Bc^1}{Binsn} \times \frac{npr^6}{Binsn}$	73	0	0	II–III instar larva
$\frac{2Bc^2}{Binsn} \times \frac{npr^6}{Binsn}$	215	0	0	Pupal

[a] Heterozygotes, flies carrying a *Br-C* mutant chromosome and a balancer chromosome (*br/Binsn*, *rbp/Binsn*, *2Bc/Binsn*, and *npr/Binsn*).

[b] Homozygous balancer, flies homozygous with respect to the balancer chromosome (*Binsn/Binsn*).

[c] Heteroallelic mutants, flies carrying different combinations of *Br-C* alleles (*br/npr*, *rbp/npr*, and *2Bc/npr*).

the formation of the dorsal appendages? To address this, heat-inducible Z1 transgenic flies (*hsp70/*Z1) were used to examine the effect of ectopic *BR-C* expression during oogenesis. Following standard heatshock (37°, 30 min) and incubation at 26° for 2–48 hr, ovaries of the *hsp70/*Z1 females were analyzed (Table 2). Flies were dissected to examine the effect of the heatshock at 2, 3, 5, 9, 24, and 48 hr after heat treatment. The first abnormalities in the egg chambers were observed 3 hr after the heatshock. The eggs laid during the first 3 hr following heatshock also have a very high hatch rate, presumably being sufficiently differentiated at the time of the heatshock for ectopic *BR-C* expression to have no effect. The results over this period do not differ significantly for the control heatshocks (Table 2). The strongest effect was observed between 4 and 6 hr. It was observed that extra dorsal appendage material was produced in the dorsal-anterior region of the eggshells (Figure 4, B–F). In most cases, dorsal appendage material appeared in the dorsal gap between the two appendages. It condensed at the base of the dorsal appendages and less material was deposited in the appendages themselves. The dorsal appendages did not elongate properly (Figure 4, B–D), presumably due to a failed migration of the follicle cells. Different phenotypes have been observed, depending on the stage of the egg chamber at the time of heatshock (Figures 4 and 5). Heatshock at the time of dorsal appendage formation, stage 11, leads to the "appendageless" phenotype or to small fused appendages (Figure 4, B and C). Appendages with abnormal shapes and/or different lengths have been observed on eggs at stages 12 and 13 at the time of heatshock (Figure 4, D–F). These observations indicate that ectopic Z1 can induce formation of ectopic dorsal appendage material. Nevertheless, the ectopic dorsal appendage material is restricted to the dorsal anterior eggshell, suggesting that the fate of the follicle cells is predetermined along the two major axes prior to the requirement for *BR-C* function in this process. It was hypothesized that the lack of *BR-C* expression in the dorsal-most follicle cells is due to high levels of expression of *pointed* (*pnt*) in those cells (Deng and Bownes 1997). The data shown here indirectly support this hypothesis. In the dorsalmost follicle cells, there could be competition between the expression of Pnt and *BR-C*. When high levels of Pnt are expressed, *BR-C* expression is inhibited in these cells. However, in heatshock lines, the *BR-C* expression would overcome the inhibition by Pnt. Thus, dorsal appendage material can be synthesized by these cells.

Although Z1 seemed to be the sole *BR-C* zinc-finger isoform expressed at high levels during oogenesis when analyzed by *in situ* hybridization, we tested Z2, Z3, and Z4 to determine if they exhibit a similar phenotype when ectopically expressed during oogenesis. Thus, *hsp70/*Z2, *hsp70/*Z3, and *hsp70/*Z4 flies were heat-

TABLE 2

Effect of heatshock on Z1–Z4 transgenic flies

Experiment	Fly stock	No. of eggs laid	Abnormal eggs (%)	Eggs hatched (%)
0–3 hr	Z1	172	5.8	94
	Z2	144	2.3	96
	Z3	134	1.5	98
	Z4	80	5.0	96
	w^{1118}	364	1.6	98
	OrR	94	2.1	99
3–5 hr	Z1	27	74	2
	Z2	68	55	8
	Z3	4	0	25
	Z4	23	43	4
	w^{1118}	137	2.5	97
	OrR	117	2.7	96
5–24 hr	Z1	10	89	6
	Z2	361	78	12
	Z3	27	42	18
	Z4	186	77	15
	w^{1118}	505	4.4	94
	OrR	363	3.6	98
24–48 hr	Z1	9	62	72
	Z2	670	33	93
	Z3	72	24	90
	Z4	390	18	97
	w^{1118}	50	1.3	100
	OrR	370	1.1	99

Z1–Z4, transgenic flies carrying constructs to misexpress the various zinc-finger isoforms of the *BR-C* in response to heatshock; w^{1118}, the host line used for the construction of the transgenic lines used as a control. OrR (wild-type) flies were also used as a control. Approximately 100 flies (50 males and 50 females) from each stock were used in the heatshock experiments.

shocked and the eggshell phenotype was examined (Table 2). It was found that ectopic dorsal appendage material is produced in the dorsal-anterior region of the eggshells by all three transgenic lines. This phenotype is similar to that exhibited by eggs of the *hsp70/Z1* flies after heatshock, suggesting that all of the four zinc-finger isoforms could be functional in dorsal appendage formation during oogenesis.

It is apparent from Table 2 that heatshock has the strongest effect on chorion morphology and egg viability in *hsp70*/Z1 flies. Z2–Z4 recovered viability to ~95% in 2 days, while Z1 recovered only to 72% during that time period. It was also found that heatshocked Z1, Z2, and Z4 flies lay abnormal eggs (Table 2). The ectopic expression of *BR-C* in Z3 flies was found to disrupt the process of egg development soon after the heatshock. We observed that some 20% of all laid eggs have aberrant micropyles, due to excess chorion formation. This could prevent the sperm entering the egg and hence subsequent development would fail due to lack of fertilization. Another possible explanation is that ectopic expression of *BR-C* can disrupt some other Bric-a-brac/Tramtrack/Broad Complex (BTB)-containing protein that can dimerize with the *BR-C* and thus modulate its function.

Heatshock alone causes eggshell defects. The data of the control experiments with the heatshocked wild-type OrR flies and w^{1118}, the host line for the transgenic flies, is presented in Table 2. We observed in the few abnormal eggs wide-branched dorsal appendages of approximately normal length. It is quite clear that the results of misexpressing *BR-C* in oogenesis significantly affects the eggshell.

Other zinc-finger isoforms are expressed in oogenesis: Although only Z1 expression was clearly observed by *in situ* hybridization we observed defects in chorion formation and morphology by overexpressing all four zinc-finger isoforms available. It became essential, therefore, to establish if this was due to some degree of functional redundancy between the isoforms with respect to eggshell development or if the other zinc fingers are, in fact, expressed at lower levels in oogenesis. To check this we used RT-PCR using primers for Z1, Z2, Z3, and Z4 and the core DNA binding domain. The organization of the zinc-finger isoforms in relation to the *BR-C* is shown in Figure 1. The primers used are illustrated in Figure 1B and should generate products of 974 bp, 780 bp, 728 bp (Z1); 320 bp (Z2); 784 bp (Z3); and 1082 bp (Z4), respectively, based on published data (Hodgetts *et al.* 1995; Bayer *et al.* 1996b). The results clearly show that all four zinc fingers are expressed in oogenesis (Figure 6). The identity of the PCR-generated products was confirmed with Southern blot analysis.

It seems likely therefore that, as in metamorphosis, all zinc-finger isoforms are expressed and function to regulate downstream gene expression. However, only Z1 is expressed at a high-enough level to detect the spatial distribution of the RNA and protein in oogenesis.

Ectopic *BR-C* expression during mid-oogenesis affects endoreplication and chorion gene amplification: Ectopic *BR-C* expression appears to induce premature production of the chorion. Figure 5, C and D, shows that the chorion is already present in stage-11 egg chamber. This could isolate the oocyte from the nurse cells and physically prevent dumping of the nurse cell components into the oocyte. This could result from an altered pattern of transcription and translation of the chorion genes, or from abnormalities in amplification of the chorion genes, or both. We investigated, therefore, whether the alterations in *BR-C* expression affected the timing or pattern of chorion gene amplification. Since the chorion is synthesized by most follicle cells, this function could be related to the earlier expression of the *BR-C*. To monitor amplification we investigated the incorporation of BrdU in the follicle cell nuclei of wild-

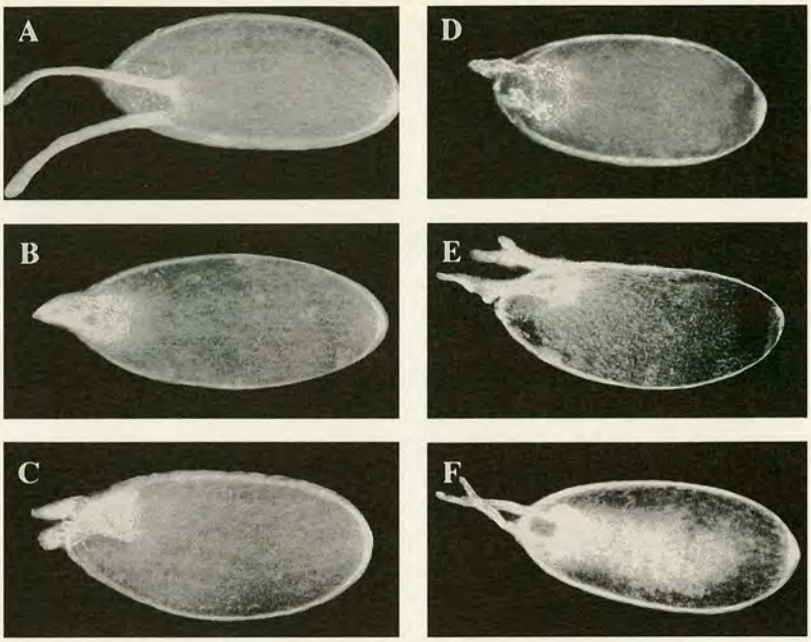


FIGURE 4.—Ectopic expression of different *BR-C* zinc-finger isoforms during oogenesis. A–F, dark-field microscopy. (A) A wild-type egg after heatshock treatment. (B and C) Two distinctive *hsp70/Z1* (527-5 and 708-1) transgenic lines exhibit a similar phenotype. Dorsal appendage materials are present in the dorsalmost region, resulting in fused, thickened, and shortened dorsal appendages. (D–F) A similar, but less severe, eggshell phenotype in *hsp70/*Z1 flies; see explanation in the text. Fused, thickened, and irregularly shaped appendages are also found in eggshells produced by the Z2, Z3, and Z4 transgenic lines.

type ovaries and in ovaries misexpressing various isoforms of the *BR-C*.

In wild-type ovaries, after eight mitotic cell divisions, the endoreplication phase of the follicle cell development of oogenesis begins (stage 6), and is completed by stage 10B; during this process the entire nucleus is labeled by BrdU. The endoreplication is asynchronous in wild-type and w^{1118} (the host strain used to produce the transgenic lines) ovaries and occurs in both nurse cells and follicle cells (Figure 7, A and C). We observed a continuous endoreplication in the nurse-cell-associated follicle cells at stage 10B (Figure 7C). This is followed by the chorion gene amplification phase when 4 spots of incorporation are seen per nucleus in the follicle cells overlying the oocyte (Figure 7, C, G, and K). These 4 spots represent amplification of the two clusters of chorion genes (Figure 7, G and K). Two are always larger, presumably due to the higher level of amplification of the cluster on chromosome 3 compared to the X-chromosome cluster (ORR-WEAVER and SPRADLING 1986; DELIDAKIS and KAFATOS 1989). This amplification was first observed at the border between the oocyte and nurse cells and it soon spread to the rest of the follicle cells. When the *BR-C* isoforms are misexpressed, there is prolonged and synchronized endoreplication until late stage 10B, followed by specific amplification of genes in each nucleus (Figure 7, D, F, H, and L). These results are observed from 3.5 to 4.5 hr after heatshock.

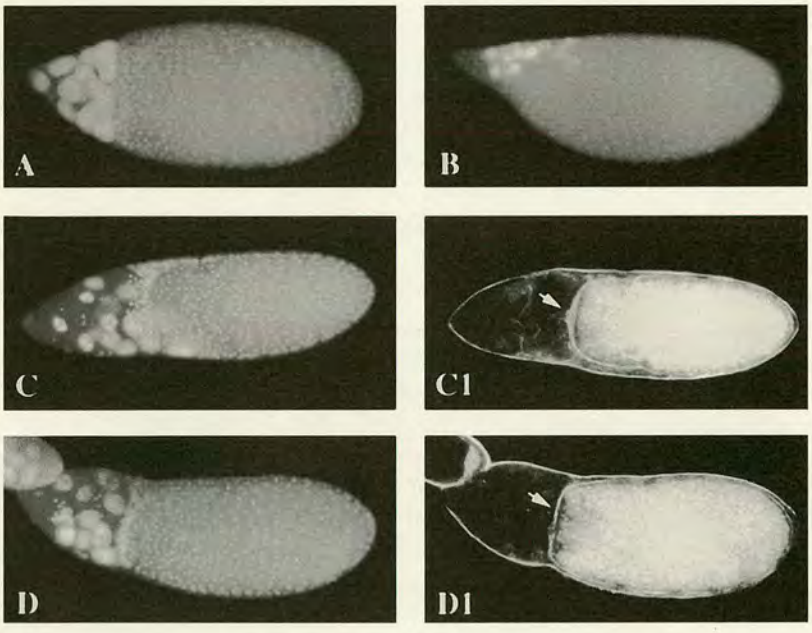


FIGURE 5.—Ectopic *BR-C* expression affects the onset of chorion synthesis. A–D, Hoechst staining; C1–D1, dark-field microscopy. (A) A wild-type egg chamber at late-stage 11. The oocyte is larger than the nurse cell complex due to the onset of nurse cell dumping into the oocyte. (B) Wild-type egg chamber, late stage 12. Dumping is complete. C–C1 (late stage 11) and D–D1 (late stage 12) show that transgenic Z1–Z4 also causes inappropriate chorion synthesis (arrows), and this blocks dumping of the nurse cell cytoplasm into the oocyte during stage 11–12 of oogenesis.

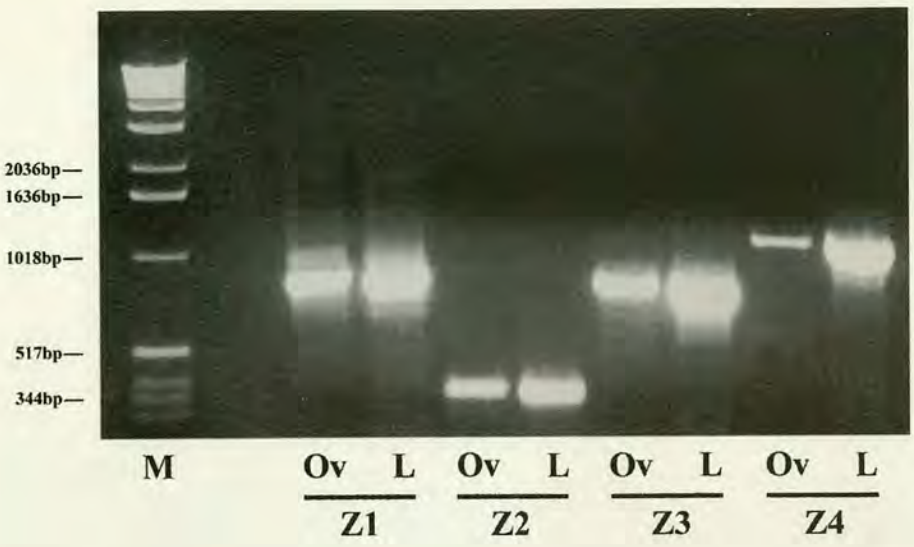


FIGURE 6.—PCR analysis of the *BR-C* transcripts in *Drosophila melanogaster* (wild-type, Oregon R) ovaries. Total RNA from ovaries was reverse transcribed and the subsequent cDNA PCR amplified using the primer pairs shown with arrows below the exon map (Figure 1B). All four PCR reactions were carried out with a common primer located at the 3′ end of the core domain of *BR-C*. The other primers are located just within, or immediately 3′ to, the respective Z1–Z4 domain. M, marker–1-kb ladder (Gibco BRL); Ov, ovaries; L, larvae (whole organism) used as a control. The primer sets generate the following products from transcribed larvae RNA: 974 bp, 780 bp, 728bp (Z1); 320bp (Z2); 784bp (Z3); 1082 bp (Z4). The PCR does not generate the 728-bp Z1 product from transcribed ovarian RNA, indicating the lack of this transcript during oogenesis.

We also observe extra spots of incorporated BrdU in the nuclei (Figure 7, H and L). There were three possible explanations for this: either the heatshock could be responsible, the homologues of the chromosomes could have separated due to a defect in the cell cycle, or there could be amplification of DNA at additional sites in the genome. The host flies used to produce the transgenic lines, w^{1118}, were heatshocked and still showed 4 spots per nucleus, so heatshock itself was not responsible for the results. We counted the number of spots per nucleus and found 6 or 12 spots in ~80% of the nuclei. Occasionally we observed up to 28 spots. If the cell cycle was affected, and the homologues had separated, we would expect to see many more than 28 spots per nucleus due to the polyploidy of the follicle cells. If the amplification sites varied we could not predict the numbers, and indeed it may well be variable. This suggests that there are other sites in the genome induced to replicate by *BR-C* overexpression. We conclude that the endoreplication of DNA and the amplification of the chorion genes depends upon the *BR-C* encoded proteins or an unknown protein that is encoded by one of the downstream targets of the *BR-C*. This observation is consistent with the report that a mutation in the *BR-C* locus causes premature arrest of chorion gene amplification (HUANG and ORR 1992).

Ectopic *BR-C* expression in relation to chorion gene expression: The chorion is produced by the columnar follicle cells to provide a shell around the egg. Later in oogenesis, two groups of cells migrate anteriorly to produce the chorionic appendages and very large amounts of chorion material. Ectopic expression of the Z1 isoforms leads to chorionic appendage deposition by extra cells lying at the anterior of the egg filling in the middorsal gap observed in wild-type eggs. Often the follicle cells fail to migrate anteriorly over the remaining nurse cells at stage 11 and they are present, therefore, at a more posterior position. The pattern of *chorion* gene expression was compared in wild-type ovaries and in those expressing the Z1 isoform ectopically. In the wild type, we observe a high concentration of chorion transcripts in all follicle cells at stage 9, prior to their translation; they then become inactive and transcripts are again seen in stages 11–14 (Figure 8, A, C, E, and G). In the ovaries with ectopic *BR-C* gene expression, examined 3.5–4.5 hr after the heatshock, the same high concentration of chorion transcripts is observed in anterior follicle cells, even though their location in relation to the nurse cells is more posterior. This suggests the expression pattern is not dependent on the *BR-C* along the anterior-posterior axis (Figure 8, B, D, F, and H). Moreover, posterior follicle cells do not produce substantially more chorion material even though they express Z1 protein after heatshock. The fact that more dorsal cells produce appendage material and express the chorion genes means that *BR-C* expression in the most dorsal anterior cells does induce additional chorion production. Thus we observed two different effects. Initially the ectopic expression of *BR-C* prevents the migration of the dorsal follicle cells in an anterior direction. Then the midline cells that normally express chorion protein at that stage start depositing chorion material in the wrong location. This results in the production of aberrant dorsal appendages. It is possible that the *BR-C* activates downstream genes which, in turn, activate the chorion genes. In anterior cells high levels of *BR-C* expression "win" over *trans*-acting repressors, but in posterior cells they do not. The *in situ* hybridization results and the observed characteristic phenotypes following ectopic *BR-C* expression are both consistent with this observation. Alternatively, *BR-C* could be essential for the cell migrations to position the follicle cells and endoreplication

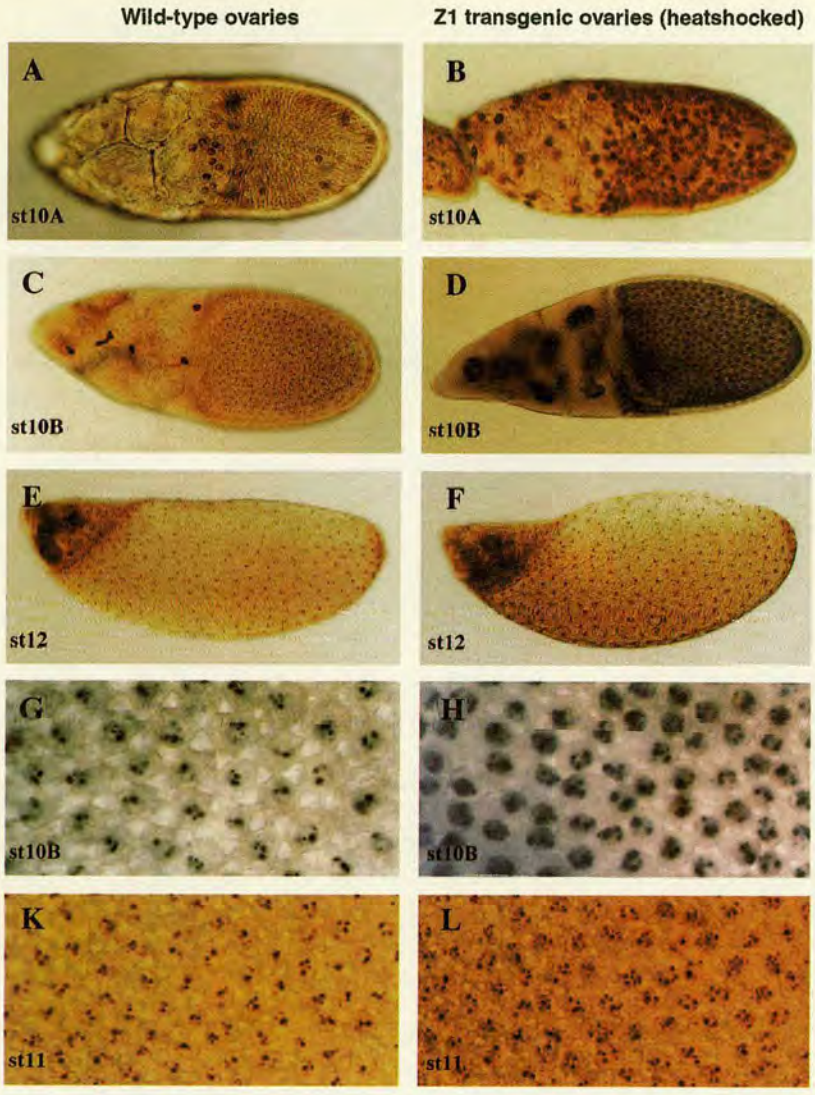


FIGURE 7.—Chorion gene amplification. BrdU incorporation associated with endoreplication and amplification. (A) BrdU incorporation in the nuclei of wild-type Drosophila (Oregon R) ovaries and w^{1118}, the host strain used for the production of the *BR-C* transgenic lines at stage 10A. The endocycles are not synchronized. Thus, just some of the nuclei are positive. (CALVI *et al.* 1998). (B) Overexpression of a *BR-C* (Z1) transgene in ovaries at stage 10A. Most of the nuclei show synchronized endoreplication. (C) Normal synchronized amplification in wild type and w^{1118} at stage 10B. By that time the endoreplication associated with the main body follicle cells is completed. (D) BrdU incorporation in the nuclei of heatshocked *BR-C* (Z1) transgenic ovaries, stage 10B. Strong BrdU incorporation is present in the nurse cells. (G) Higher magnification of wild-type and w^{1118} nuclei at stage 10B to show the four spots of amplification normally occurring in wild-type ovaries. (H) BrdU incorporation in the nuclei of *BR-C* (Z1) expressing ovaries 3 hr after the heatshock, stage 10B. Some of the nuclei contain extra spots of replication. The amplification pattern overlaps with the labeling due to continued endoreplication. (E–K) The nuclei of heatshocked w^{1118} ovaries still show four spots of amplification. (F–L) The nuclei of *BR-C* (Z1) transgenic ovaries 4.5 hr after heatshock exhibit a multispotted pattern.

of the chorion genes but not in regulating chorion gene expression.

DISCUSSION

The *BR-C* complementing groups and zinc-finger isoforms: It has been shown by genetic analysis that rbp^+ function is required for dorsal appendage formation, and it was observed that Z1 is the sole zinc-finger isoform expressed at high levels in the appendage-producing cells. These observations are compatible with the report that Z1 provides rbp^+ function (BAYER *et al.* 1997). It was also found that rbp^+ function is partially provided by Z4, but is not provided by Z2 and Z3 (BAYER *et al.* 1997). However, heatshock-induced expression of all four zinc-finger isoforms (Z1, Z2, Z3, and Z4) leads to a similar phenotype of extra dorsal appendage material production in the dorsal gap, indicating that the different transcripts may substitute for each other functionally in dorsal appendage formation.

The homozygous viable mutant br^1 was the first mutant identified in the *BR-C* locus. It exhibits a broad wing phenotype and fails to complement other mutations that are categorized in the *br* complementation group. However, the complementation analysis presented in this article suggests that the br^1 mutations also partially remove *rbp* function. Therefore, it is in fact a *2Bab* allele. It is known the *2Bab* mutations cause reduction of both Z1 and Z2 expression. Thus in the br^1/br^{A47} and br^1/npr^6 females, both Z1 and Z2 are reduced. The reduction of Z1 levels results in the reduction of dorsal appendages, while no effect is produced by the reduction of Z2 levels. This is why no defects were observed in eggs produced by br^1/br^5 and br^1/br^6 females.

The mutant phenotypes clearly show the need for the *BR-C* in chorionic appendage formation. PCR experiments have shown that all zinc-finger isoforms are in fact expressed in oogenesis, but as yet we have no evidence that they perform different functions. Neither do we know the spatial and temporal distribution of Z2–Z4, which are not present at sufficiently high levels for detection by *in situ* hybridization. Overexpression

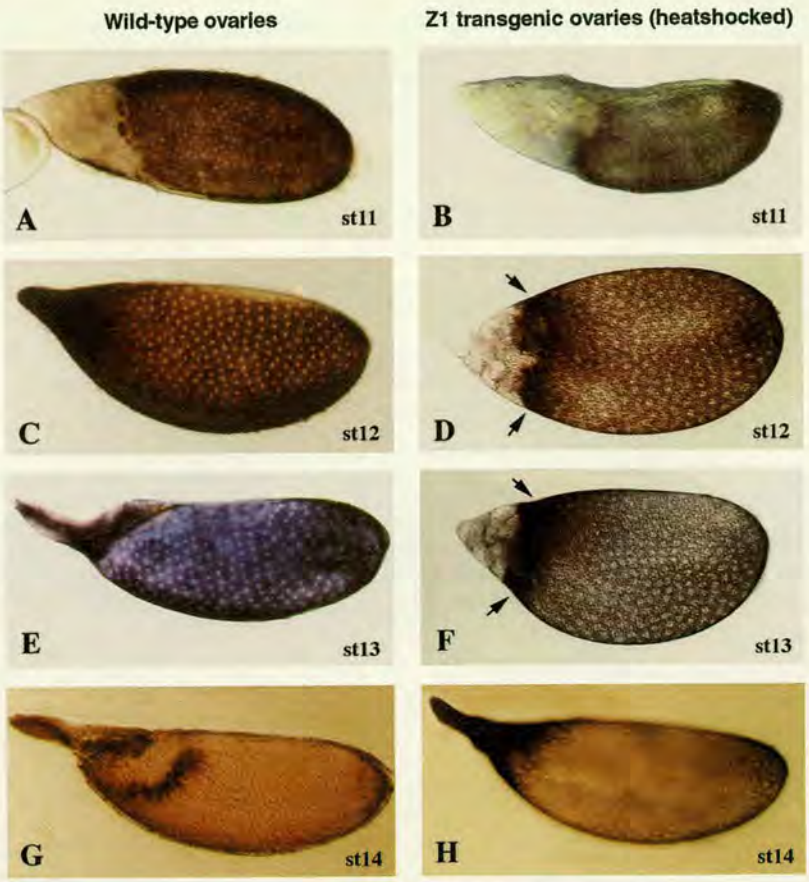


FIGURE 8.—*In situ* hybridization to chorion RNA in ovaries. (A, C, E, G) *In situ* hybridization to wild-type (Oregon R) Drosophila ovaries with a DIG-labeled probe hybridizing to chorion mRNA. The chorion probe, 7C8, was generated from cDNA encoding the c38 chorion protein, which maps to the X chromosome. Similar results were obtained using the alternative 7B7 probe (cDNA encoding c15 protein), which maps to a site on the third chromosome. For more information about the probes see SPRADLING *et al.* (1980). (B) *In situ* hybridization to heatshocked Drosophila *BR-C* (Z1) transgenic ovaries with 7C8 probe, stage 11. At that stage an increase in the amount of chorion mRNA is observed at the border between the nurse cells and the oocyte. The misexpression of the *BR-C* leads to abnormal migration of the anterior follicle cells. (D) *In situ* pattern of chorion mRNA expression in heatshocked *BR-C* (Z1) transgenic ovaries, stage 12. A gap (arrows) is formed at the anterior pole as a result of failed migration of the follicle cells. (F) Strong expression of the chorion gene clusters in heatshocked *BR-C* (Z1) transgenic ovaries at stage 13. The variations in the egg chamber shape result from the heatshock-induced misexpression of *BR-C*. (H) Strong continuous expression of chorion mRNA in heatshocked *BR-C* (Z1) transgenic ovaries, stage 14. The anterior gap is not present because the nurse cells have completed dumping and the follicle cells have migrated correctly by the time of heatshock.

studies using transgenic flies carrying heatshock-controlled Z1 and Z4 isoforms lead a failure of proper migration of the follicle cells that will secrete the appendages, premature chorion deposition, and abnormal appendage formation.

Relationship between the *BR-C* and chorion production: We have shown here that the *BR-C* is important for controlled DNA replication in oogenesis. Overexpression does not affect the timing of the onset of endoreplication and amplification, but endoreplication is prolonged beyond that observed in wild-type ovaries and it leads to additional replication sites in the genome. These additional sites presumably share sequence similarities with the *cis*-acting sites regulating chorion gene amplification. This suggests that the *BR-C* is a key regulator of endoreplication and chorion gene amplification. The early *BR-C* expression pattern is in all the follicle cells and it is presumably at this stage that it is involved in this function. The expression of the *BR-C* is first observed in wild-type flies at stage 6 and it is also at stage 6 that the endoreplication cycles begin. Since we did not observe premature endoreplication with *BR-C* overexpression, presumably other components essential for endoreplication are absent until stage 6 of oogenesis. The active role of the *BR-C* in endoreplication is also apparent from the fact that we observed prolonged incorporation of BrdU in the nurse cell nuclei when the *BR-C* is overexpressed. This presumably results in expression of the *BR-C* in the nurse cells, where it is normally not expressed. This shows that the proteins encoded by the *BR-C* can function to prolong replication of DNA even in cells where it is not normally used to control this process. It also suggests that an alternative regulator for DNA replication to the *BR-C* is used in the nurse cells. In normal development, the later *BR-C* expression, which is maintained in the dorsal-anterior cells making the appendages, is probably needed for cell migration, chorion deposition, and other follicle cell differentiation events.

It is possible that the initial activation of the *BR-C* in all follicle cells is regulated by the ecdysone/USP heterodimer (YAO *et al.* 1992; HORNER *et al.* 1995) as is observed in metamorphosis. There is also some evidence that in intermoult puffs a second heterodimer with DHR38 (a nuclear receptor related to NGF1-B from mammals) can compete for the same binding sites (CRISPI *et al.* 1998), so it is also possible that other DNA binding proteins, besides the ecdysone receptor, will be involved in *BR-C* expression. This would suggest a role for ecdysone during these stages of oogenesis. Later, the *BR-C* could be repressed in specific follicle cells by competitive and cooperative interactions with other gene products initiated by the *grk* and *dpp* signaling pathways. There is a significant amount of evidence that ecdysone and juvenile hormone [which binds to ultra spiracle proteins (USP)] are important for the progress

of oogenesis (WILSON 1982; BOWNES 1989, 1994) and we have recently shown that there is a control point in oogenesis that regulates whether egg chambers will proceed with development or undergo apoptosis, which is regulated by the balance of juvenile hormone and ecdysone (SOLLER *et al.* 1999). However, there is little evidence as to precisely what role these hormones have in regulating oocyte development and egg chamber differentiation. We have shown that the ecdysone receptor is present in the follicle cells at the time *BR-C* is activated (D. MAUCHLINE, W.-M. DENG and M. BOWNES, unpublished results) by antibody staining.

Once activated, as we have shown, the *BR-C* gene is involved in endoreplication, the selective amplification of the chorion genes, and in the subsequent morphogenesis of the chorionic appendages. CALVI *et al.* (1998) have recently shown that the selective amplification of the chorion genes is closely linked with the cell cycle and the cycles of endoreplication that occur in the follicle cells earlier. Somehow the chorion genes escape the rereplication controls that influence other parts of the genome. Our BrdU labeling experiments confirm their results on the timing of endoreplication and chorion amplification and the close association between endoreplication and selective amplification. Using overexpression of the *BR-C* we see not only the two extra sites they mentioned that may represent another chorion gene amplified for a function in later oogenesis, but also a number of additional sites. These may be sites with sequence similarity to the *cis*-acting sites regulating amplification. CALVI *et al.* (1998) propose that there are amplification complexes located at chorion genes. Whether the *BR-C* encoded proteins are associated with these complexes or regulate the synthesis of one or more of their components remains to be elucidated. We have confirmed the link between endoreplication and chorion amplification and shown that it involves the *BR-C*. This may therefore provide a crucial link between hormones and the control of the cell cycle, and hence of differentiation, of the egg chamber during oogenesis.

In summary, our working model would be that the *BR-C* is activated by ecdysone in all follicle cells at stage 6 of oogenesis where its key function is the control of endoreplication, and then selective amplification. Later, when it is turned off in all but the anterior-dorsal follicle cells that will secrete the appendages it has a second set of functions and is involved in the migration of cells and morphogenesis of the chorionic appendages. Recently this link between ecdysone, the *BR-C*, and morphogenesis has also been described for the progression of the furrow in the developing eye imaginal disc of Drosophila (BRENNAN *et al.* 1998).

We thank C. Bayer and J. Fristrom for sending us heatshock *BR-C* flies and *BR-C* alleles, L. Restifo for *rpl*^{2} fly stocks, Gregory Guild for sending anti-*BR-C* core and Z1 and Z3 antibodies, Alan Spradling for chorion gene probes, and Brian Calvi for help with developing the BrdU labeling techniques. We also thank D. Zhao and other colleagues in M. Bownes' lab for helpful discussions and Sheila Milne for typing the manuscript. W.-M.D. and G.T. were supported by Darwin Trust scholarships. This research was supported by the Wellcome Trust.

LITERATURE CITED

ASHBURNER, M., 1974 Sequential gene activation by ecdysone in polytene chromosomes of *Drosophila melanogaster*. II. The effects of inhibitors of protein synthesis. Dev. Biol. **39:** 141–157.

BARDWELL, V. J., and R. TREISMAN, 1994 The POZ domain: a conserved protein-protein interaction motif. Genes Dev. **8:** 1664–1677.

BAYER, C., L. VON KALM and J. W. FRISTROM, 1996a Gene regulation in imaginal disc and salivary gland development during Drosophila metamorphosis, pp. 321–361 in *Metamorphosis*. Academic Press, San Diego.

BAYER, C. A., B. HOLLEY, and J. W. FRISTROM, 1996b A switch in broad-complex zinc-finger isoform expression is regulated posttranscriptionally during the metamorphosis of Drosophila imaginal discs. Dev. Biol. **177:** 1–14.

BAYER, C. A., L. VON KALM and J. W. FRISTROM, 1997 Relationships between protein isoforms and genetic functions demonstrate functional redundancy at *Broad-Complex* during Drosophila metamorphosis. Dev. Biol. **187:** 267–282.

BELYAEVA, E. S., M. G. AIZENZON, V. F. SEMESHIN, I. KISS, K. KOCZYA *et al.*, 1980 Cytogenetic analysis of the 2B3/4-2B11 region of the X-chromosome of *Drosophila melanogaster*. I. Cytology of the region and mutant complementation groups. Chromosoma **81:** 281–306.

BOWNES, M., 1989 The roles of juvenile hormone ecdysone and the ovary in the control of Drosophila vitallogenesis. J. Insect Physiol. **35:** 409–414.

BOWNES, M., 1994 The regulation of the yolk protein genes, a family of sex differentiation genes in *Drosophila melanogaster*. Bioessays **16:** 745–752.

BRENNAN, C. A., M. ASHBURNER and K. MOSES, 1998 Ecdysone pathway is required for furrow progression in the developing Drosophila eye. Development **125:** 2653–2664.

CALVI, B. R., M. A. LILLY and A. C. SPRADLING, 1998 Cell cycle control of chorion gene amplification. Genes Dev. **12:** 734–744.

CHAO, A. T., and G. M. GUILD, 1986 Molecular analysis of the ecdysterone-inducible 2B5 'early' puff in *Drosophila melanogaster*. EMBO J. **5:** 143–150.

CRISPI, S., E. GIORDANO, P. P. D'AVINO and M. FURIA, 1998 Cross-talking among Drosophila nuclear receptors at the promiscuous response element of the *ng-1* and *ng-2* intermolt genes. J. Mol. Biol. **275:** 561–574.

DELIDAKIS, C., and F. C. KAFATOS, 1989 Amplification enhancers and replication origins in the autosomal chorion gene cluster of Drosophila. EMBO J. **8:** 891–901.

DENG, W.-M., and M. BOWNES, 1997 Two signalling pathways specify localised expression of the *Broad-Complex* in Drosophila eggshell patterning and morphogenesis. Development **124:** 4639–4647.

DIBELLO, P. R., D. A. WITHERS, J. W. FRISTROM and G. M. GUILD, 1991 The Drosophila *Broad-Complex* encodes a family of related proteins containing zinc fingers. Genetics **129:** 385–397.

EMERY, I. F., V. BEDIAN and G. M. GUILD, 1994 Differential expression of *Broad-Complex* transcription factors may forecast tissue-specific developmental fates during Drosophila metamorphosis. Development **120:** 3275–3287.

HODGETTS, R. B., W. C. CLARK, S. L. O'KEEFE, M. SCHOULS, K. CROSSGROVE *et al.*, 1995 Hormonal induction of Dopa decarboxylase in the epidermis of Drosophila is mediated by the *Broad-Complex*. Development **121:** 3913–3922.

HORNER, M. A., T. CHEN and C. S. THUMMEL, 1995 Ecdysteroid regulation and DNA binding properties of Drosophila nuclear hormone receptor superfamily members. Dev. Biol. **168:** 490–502.

HUANG, R. Y., and W. C. ORR, 1992 *Broad-Complex* function during oogenesis in *Drosophila melanogaster*. Dev. Genet. **13:** 277–288.

KARIM, F. D., G. M. GUILD and C. S. THUMMEL, 1993 The Drosophila *Broad-Complex* plays a key role in controlling ecdysone-regulated gene expression at the onset of metamorphosis. Development **118:** 977–988.

KISS, I., A. H. BEATON, J. TARDIFF, D. FRISTROM and J. W. FRISTROM,

1988 Interactions and developmental effects of mutations in the *Broad-Complex* of *Drosophila melanogaster*. Genetics **118:** 247–259.

LILLY, M. A., and A. C. SPRADLING, 1996 The Drosophila endocycle is controlled by Cyclin E and lacks a checkpoint ensuring S-phase completion. Genes Dev. **10:** 2514–2526.

LIN, H., and A. C. SPRADLING, 1993 Germline stem cell division and egg chamber development in transplanted Drosophila germaria. Dev. Biol. **159:** 140–152.

ORR, W. C., V. K. GALANOPOULOS, C. P. ROMANO and F. C. KAFATOS, 1989 A female sterile screen of the *Drosophila melanogaster* X chromosome using hybrid disgenesis: identification and characterization of egg morphology mutants. Genetics **122:** 847–858.

ORR-WEAVER, T. L., 1991 Drosophila chorion genes: cracking the eggshell's secretes. Bioessays **13:** 97–105.

ORR-WEAVER, T. L., and A. C. SPRADLING, 1986 Drosophila chorion gene amplification requires an upstream region regulating s 18 transcription. Mol. Cell. Biol. **6:** 4624–4633.

SOLLER, M., M. BOWNES and E. KUBLI, 1999 Control of oocyte maturation in Drosophila. Dev. Biol. **208:** 337–351.

SPRADLING, A. C., 1993 Developmental genetics of oogenesis, pp. 1–70 in *The Development of Drosophila melanogaster*, Vol. 1, edited by M. BATE and A. M. ARIAS. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

SPRADLING, A. C., and A. P. MAHOWALD, 1980 Amplification of genes for chorion proteins during oogenesis in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **77:** 1096–1100.

SPRADLING, A. C., M. E. DIGAN and A. P. MAHOWALD, 1980 Two clusters of genes for major chorion proteins of *Drosophila melanogaster*. Cell **19:** 905–914.

TAUTZ, D., and C. PFEIFLE, 1998 A nonradioactive in-situ hybridisation method for the localisation of specific RNAs in Drosophila embryos reveals a translational control of the segmentation gene *hunchback*. Chromosoma **98:** 81–85.

WILSON, T. G., 1982 A correlation between juvenile hormone deficiency and vitellogenic oocyte degeneration in *Drosophila melanogaster*. Wilhem Roux's Arch. Entw. Mec. Org. **1919:** 257–263.

YAO, T-P., W. A. SEGRAVES, A. E. ORO, M. MCKEOWN and R. M. EVANS, 1992 Drosophila ultraspiracle modulates ecdysone receptor function via heterodimer formation. Cell **71:** 63–72.

ZHIMULEV, I. F., E. S. BELYAEVA, O. M. MAZINA and M. L. BALASOV, 1995 Structure and expression of the *br-c* locus in *Drosophila melanogaster* (Diptera: Drosophilidae). Eur. J. Entomol. **92:** 263–270.

ZOLLMAN, S., D. GODT, G. G. PRIVE, J. L. COUDERC and F. A. LASKI, 1994 The BTB domain, found primarily in zinc-finger proteins, defines an evolutionarily conserved family that includes several developmentally regulated genes in Drosophila. Proc. Natl. Acad. Sci. USA **91:** 10717–10721.

Communicating editor: T. SCHÜPBACH

Identification and Phylogenetic Analysis of *Drosophila melanogaster* Myosins

George Tzolovsky, Hadas Millo, Stephen Pathirana, Timothy Wood, and Mary Bownes
Institute of Cell and Molecular Biology, University of Edinburgh

Myosins constitute a superfamily of motor proteins that convert energy from ATP hydrolysis into mechanical movement along the actin filaments. Phylogenetic analysis currently places myosins into 17 classes based on class-specific features of their conserved motor domain. Traditionally, the myosins have been divided into two classes depending on whether they form monomers or dimers. The conventional myosin of muscle and nonmuscle cells forms class II myosins. They are complex molecules of four light chains bound to two heavy chains that form bipolar filaments via interactions between their coiled-coil tails (type II). Class I myosins are smaller monomeric myosins referred to as unconventional myosins. Now, at least 15 other classes of unconventional myosins are known. How many myosins are needed to ensure the proper development and function of eukaryotic organisms? Thus far, three types of myosins were found in budding yeast, six in the nematode *Caenorhabditis elegans*, and at least 12 in human. Here, we report on the identification and classification of *Drosophila melanogaster* myosins. Analysis of the *Drosophila* genome sequence identified 13 myosin genes. Phylogenetic analysis based on the sequence comparison of the myosin motor domains, as well as the presence of the class-specific domains, suggests that *Drosophila* myosins can be divided into nine major classes. Myosins belonging to previously described classes I, II, III, V, VI, and VII are present. Molecular and phylogenetic analysis indicates that the fruitfly genome contains at least five new myosins. Three of them fall into previously described myosin classes I, VII, and XV. Another myosin is a homolog of the mouse and human PDZ-containing myosins, forming the recently defined class XVIII myosins. PDZ domains are named after the postsynaptic density, disc-large, ZO-1 proteins in which they were first described. The fifth myosin shows a unique domain composition and a low homology to any of the existing classes. We propose that this is classified when similar myosins are identified in other species.

Introduction

The past decade has seen a significant increase in research on myosins. A major effort has been put into finding novel members of this family of actin-based motor proteins. More than 16 classes of myosins have been discovered and characterized, and this number is still rising (Hodge and Cope 2000; Sellers 2000). These myosins are often referred to as unconventional (Mooseker and Cheney 1995). The total number of known myosins is 17 if the conventional two-headed filament forming myosin II is included in the classification. Myosins have been identified in a wide variety of eukaryotic organisms. Some myosin classes are found in phylogenetically diverse organisms, whereas others, which have arisen later in evolution, have been found in only a single organism (Hodge and Cope 2000).

Current research concentrates on the functional analysis of these new types of myosins. A number of studies suggest that these motors play important roles in a variety of cellular functions, including organelle, RNA and protein transport, maintenance of the cell architecture, cell movements, and signal transduction (table 1).

All known myosins comprise an N-terminal head domain, a neck regulatory domain, and a specific carboxy-terminal tail domain (fig. 1) (Mooseker and Cheney 1995). The head or motor domain contains ATP- and actin-binding sites and is responsible for the mechanochemical properties of the protein (Gilbert and Mackey 2000). Myosins show an actin-stimulated Mg^{2+} ATPase activity, thus converting the energy stored in ATP into mechanical force (Volkmann and Hanein 2000). The latter is used to move the myosin molecules along the actin filaments or to translocate other molecules (Hasson and Mooseker 1995; Langford 1995).

The neck domain contains regulatory sites, composed of IQ (isoleucine-glutamine) motifs, repeats of 23–30 aa (Mercer et al. 1991; Rhoads and Friedberg 1997). Each IQ motif provides a binding site for a calmodulin or a related protein of the EF-hand family (Kawasaki, Nakayma, and Kretsinger 1998). EF proteins have helix-loop-helix motifs in which the loop contains highly conserved residues that bind Ca^{2+} ions. The size of the neck domain varies from one to seven IQ tandem repeats. In addition, the neck is often the site of alternative splicing. This produces necks with variable lengths (variable number of IQ repeats), which are associated with the regulatory function. In general, calmodulin activates a diverse group of target cellular proteins when bound to Ca^{2+}. Interestingly, most of the unconventional myosins carry IQ motifs that bind calmodulin with higher affinity in the absence of Ca^{2+}.

After the neck domain, each myosin has a highly divergent tail domain. A subset of myosin tails has predicted coil-coil α-helical domains, which promote the formation of dimmers, a typical example being the two-headed conventional myosin II. Some other myosins lack coiled-coil domains but contain structural domains found in other proteins (table 1).

The classification of myosins is based on the sequence comparison of their core motor domains (myosin head), equivalent to amino acids 88–780 of chicken skeletal myosin II (Cope et al. 1996). The motor domain is highly conserved among all myosins, reflecting the high conservation of its function. However, they have a number of class-specific features (characteristic inserts

Key words: unconventional myosins, evolution, genome project.

Address for correspondence and reprints: Mary Bownes, Institute of Cell and Molecular Biology, University of Edinburgh, King's Buildings, Edinburgh EH9 3JR, United Kingdom.
E-mail: mary.bownes@ed.ac.uk.

Mol. Biol. Evol. 19(7):1041–1052. 2002

Table 1
Domain Structure and Function of Myosin Classes

Class Myosins	Number of IQ Motifs	Number of Heavy Chains	Other (N- and C-terminal) Domains	Function
I				
Subclass 1	One or two IQ	1 (monomer)	TH1, GPA or GPQ, and SH3	Vesicle transport, cell growth, and cell motility
Subclass 2	Three to six IQ	1 (monomer)	TH1	Function in the microvilli of the brush border
Subclass 3	Three IQ	1 (monomer)	TH1	Epithelial morphogenesis and hair cells function
Subclass 4	Two IQ	1 (monomer)	TH1	Epithelial morphogenesis
II				
Muscle	Two IQ	2 (dimer)	CC	Smooth or skeletal muscle contraction
Nonmuscle	Two IQ	2 (dimer)	CC	Maintainance of the cell architecture, cell motility, and phagocytosis
III	One or two IQ	1 (monomer)	N-terminal protein kinase	Role in prototransduction
IV	One IQ	1 (monomer)	MyTH4 and SH3	? (Only in *Acanthamoeba* species)
V	Six IQ	2 (dimer)	CC, transmembrane, specific C-terminal globular domain	Membrane trafficking, polarized cell growth, vesicle, protein and/or mRNA transport
VI	One IQ	2 (dimer) ?	CC, reverse gear, specific C-terminal globular domain	Vesicle transport, epithelial morphogenesis, and stereocilia function. Spermatid individualization. Moves toward the "-" end of actin filaments
VII	Four or five IQ	2 (dimer) ?	CC, MyTH4, FERM, and SH3	Membrane trafficking, hair and photoreceptor cells function
VIII	Three or four IQ	2 (dimer) ?	CC, serine rich domain, specific C-terminal domain	Cell wall function in plants, intracellular transport
IX	Four to six IQ	1 (monomer)	N-terminal extension, zinc binding, and Rho-GAP domains	Signaling (GTPase activating)
X	Three IQ	2 (dimer) ?	CC, PH, MyTH4, FERM	Localized to regions of dynamic actin. Signal transduction
XI	Five to six IQ	2 (dimer) ?	CC	Vesicular transport in plants
XII	IQ ?	2 (dimer) ?	CC, N-terminal extension, and MyTH4 domains	? (In *C. elegans* only)
XIII	Four to seven IQ	1 (monomer)		? (In plants only).
XIV	IQ ?	1 (monomer)		? (In *Toxoplasma* and *Plasmodium* species)
XV	Two to three IQ	1 (monomer)	N-terminal extension MyTH4, FERM, and SH3	Hair cell function
XVI	Two IQ	1 (monomer)	Ankyrin repeats	Neuronal cell migration
XVII	IQ ?	1 (monomer)	Chitin synthase domain	? (In *Pyricularia* and *Emiricella* species)
XVIII	One or two IQ	2 (dimer) ?	CC, KE, and PDZ domain	? Maintenance of the stromal cell architecture

NOTE.—The number of heavy chains reflects their ability to dimerize (based on coiled-coil predictions). Key to domain abbreviations: TH1, Tail Homology Basic 1 domain; GPA, glycine-proline-alanine–rich domain; GPQ, glycine-proline-glutamine–rich domain; SH3, Tail Homology 3 domain (binds to proline-rich motifs); CC, coiled-coil; MyTH4, Myosin Tail Homology 4 domain; Rho-GAP domain (activates small GTPases of the Rho family); PH, Pleckstrin Homology domain; KE, lysine-glutamate–rich domain; PDZ or DHR (Dlg homologous region). For recent reviews on the structure and properties of the unconventional myosins see Wu, Jung, and Hammer (2000), Oliver, Berg, and Cheney (1999), Baker and Titus (1998), and Mermall, Post, and Mooseker (1998). In the construction of this table, data were used from these reviews, the myosin home page, and Cope et al. (1996).

or substitutions), which might be important in defining the precise function of a given myosin. For further information see the Myosin home page at http://www.mrc-lmb.cam.ac.uk/myosin/myosin.html (Hodge and Cope 2000). Phylogenetic analysis of the tail domain sequences produces similar results, indicating that heads and tails have coevolved (Korn 2000).

Five myosin genes have been identified in yeast (*Saccharomyces cerevisiae*), falling into three classes: two class I myosins, one class II myosin, and two class V myosins (Brown 1997). It was suggested that the whole yeast genome had undergone a duplication in ancient times, followed by a number of modifications. As a result, a small fraction of the genes were retained in duplicate (most of them being deleted), thus explaining the loss of the second myosin II gene (Wolfe and Shields 1997). *Saccharomyces cerevisiae* is the organism with the lowest known number of myosin genes. This demonstrates that a eukaryote can function with a set of only three types of myosins.

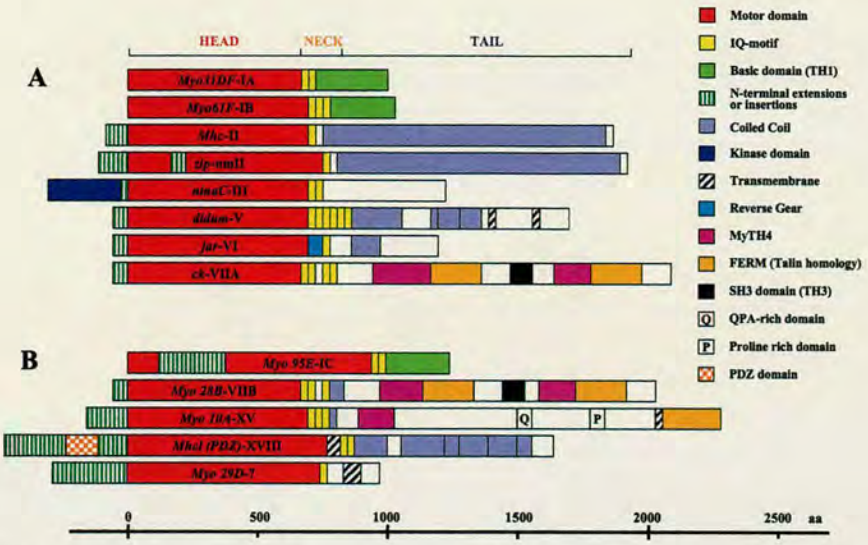


FIG. 1.—Schematic diagram of the domain structure of *Drosophila* myosins. (*A*) The previously described *Drosophila* myosins. Their sequence and structure have been studied and confirmed. For DNA and protein accession numbers see table 2. (*B*) The five new or previously uncharacterized *Drosophila* myosins. Their domain structure is based on the predicted cDNAs for these myosins (see table 2). The presence of some domains (e.g., the PDZ domain in Mhcl myosin) was verified with RT-PCR. A color key to the domain names is given on the left. Some of the names are abbreviated: MyTH4, Myosin Tail Homology 4 domain; FERM, Band 4.1, ezrin, radixin, moesin-homology domain; QPA, glutamine-proline-alanine–rich domain; PDZ, a domain with potential for targeting membrane sites.

Multicellular organisms have the ability to express some 10–40 myosin genes encoding at least six types of myosins. It seems that multicellular organisms require a multitude of specialized myosins. This has raised the question of what is the degree of functional redundancy between the classes and between the members of a given myosin class.

So far, 11 myosin genes have been identified in the slime mould *Dictyostelium discoideum*. Despite the fact that this is one of the simplest multicellular organisms, it expresses a diverse set of myosin genes (Soldati, Geissler, and Schwarz 1999). They encode at least six different classes of highly specific myosins. There are six class I myosins (MyoA, B, C, D, E, K, and probably MyoF) and one member each of class II (MhcA), class VII (MyoI) (Titus 1999), and class XI (MyoJ) (Hammer and Jung 1996). The highly divergent MyoM is still to be classified (Schwarz, Geissler, and Soldati 1999).

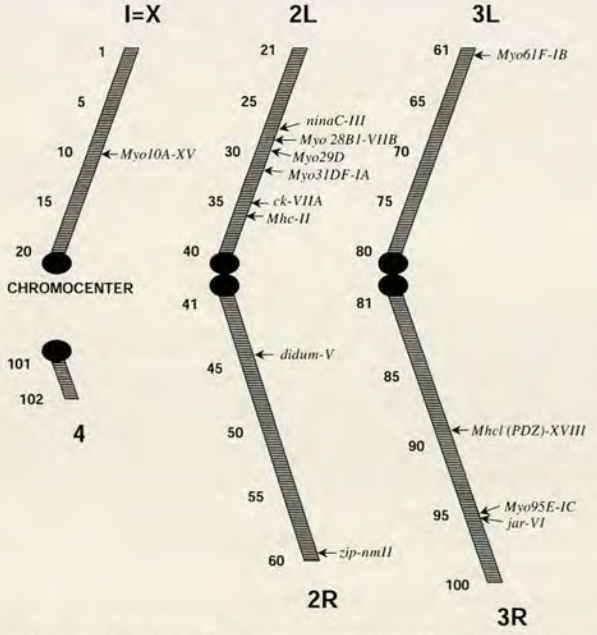


FIG. 2.—Polytene chromosome position of *D. melanogaster* myosin genes (based on the data provided by the Berkeley Drosophila Genome Project). The figure presents all myosin genes, new and previously identified, in relation to their chromosomal locations. So far, no myosin genes have been found on the fourth chromosome.

Fourteen myosin genes have been identified in the nematode *Caenorhabditis elegans* (Baker and Titus 1997). They encode two structurally distinct class I, six class II, one class V, two class VI, one class VII, and one class IX myosins. It was found that *C. elegans* has a highly divergent type of myosin, which is the founding and only member of class XII myosins.

The situation with vertebrates appears even more complex. They express some 40 myosin genes grouped into 12 classes. In humans, there are 8 class I, 16 class II, 2 class III, 3 class V, 1 class VI, 2 class VII, 2 class IX, 1 class X, 2 class XV, and 1 class XVI myosins (Hasson et al. 1996; Berg et al. 2000; Berg, Powell, and Cheney 2001). Recent studies have discovered two PDZ-containing myosins (Furusawa et al. 2000), as well as a novel unclassified myosin (Berg, Powell, and Cheney 2001). PDZ domains are named after the postsynaptic, disc-large, ZO-1 proteins in which they were first described.

In *Drosophila* eight different myosin genes have been described thus far (figs. 1*A*, 2, and table 2). There are two class I myosins, members of subclass 3 (myosin IB) and subclass 4 (myosin IA) (Morgan et al. 1994; Mooseker and Cheney 1995; Morgan, Heinzelman, and Mooseker 1995). Only a single muscle myosin II gene was found in *Drosophila* (Hastings and Emerson 1991; Bernstein and Milligan 1997). It encodes more than 13 protein isoforms with complex temporal and spatial ex-

Table 2
A List of *Drosophila melanogaster* Myosins

Gene–Protein	Class	Chromosome Position	Genomic DNA Length (bp)	mRNA Length (bp)	Protein Length (aa)	EST
Myo31DF (CG7438)* Myosin IA	I	2L (31E3-6)	AE003628 (15,084)	U07595 (3,886)	AAA19590, AAF52966 (1,011)	LP03284.5, LP04189.5
Myo61F (CG9155)* Myosin IB	I	3L (61F6-7)	AE003471, AC005847 (5,057)	U07596 (3,382)	AAA19591 (1,026)	GH13657.5, GH13670.5, GH25605.5, LP08690.5, LP05759.5, CK01057.5, GH23689.5
Myo95E (CG5501) Myosin IC	I	3R (95E1-4)	AE003746 (4,604)	AF454350 (5,285), AF454351 (5,225), AF454352 (5,301)	AF454350 (1,278), AF454351 (1,258)	GH25580.5, GH25580.3, HL05373.5
zip (zipper) (CG15792)* Nonmuscle myosin-II	II	2R (60E7-8)	AE003465 (21,711)	U35816 (6,734)	AAF47311, AAB09048 (2,056)	GM04963.5, SD07905.5, GH27070.5
Myo (CG17927)* Myosin heavy chain	II	2L (36A8-9)	M61229, AE003652, AC005119 (22,663)	M61229 (5,889)	AAA28686, AAA28687 (1,962)	GH21445.5, GH06021.5, GH02291.5, LP07131.3, LD16079.5, HL04385.5, GH11211.5, LD35610.3, HL03720.5, GM03715.5, GM03876.5
ninaC (CG5125)* (Nina C protein)	III	2L (27F5-6)	AE003617, J03131 (7,220)	M20230 (4,809), M20230 (3,643)	AAA28718 (1,501), AAA28719 (1,135)	LP02603.5, LP02603.3
didum (CG2146)* Myosin V	V	2R (43D1-3)	AC004280, AE003841 (8,217)	AF003826 (6,214)	AAF59241, AC99496 (1,792)	GH04445.5, GH04445.3
jar (jaguar) (CG5695)* Myosin VI	VI	3R (95F1-2)	AE003747 (6,312)	X67077 (4,280)	AAF56269, CAA47462 (1,253)	HL03149.5, GH09735.5
ck (crinkled) (CG7595)* Myosin VIIA	VII	2L (35C1)	AE003646 (12,392)	AE003646 (7,030)	AAF53435, AAF44915 (2,167)	LD14917.5, LD14917.3
Myo28B1 (CG6976) Myosin VIIB	VII	2L (28B3-C1)	AE003618, AC005834 (10,519)	AF233269 (6,590)	AAF52536, AAF34810 (2,129)	GH25551.5, GH25551.3
Myo10A (CG2174) Myosin XV	XV	X (10A1)	AE003484 (10,235)	AE003484, AF454346 (7,275)	AAF47980, Q9VZ48, (2,424) (2,333)	LP03318.5
Mhcl (CG10218) Myosin heavy chain–like (PDZ) Myosin	XVIII	3R (89B7)	AE003711 (21,620)	AE003711, AF454347, AY051503	AAK92927, AF454347, AAF55271, AAF55272 (6,603) (6,603) (6,420)	GH03004.5, GM10420.5, GH04935.5, LP04491.5, GH15471.5, GH15471.3
Myo29D (CG10595) Myosin	?	2L (29D1)	AE003621 (5,732)	AF454348 (4,402), AF454349 (2,530)	AF454348 (1,313), AF454349 (689), AAF52683, Q9VLK6, AF405293, AAK97502	LD47348.5, LD47348.3, LP07160.5, AI124339

NOTE.—Myosins denoted with * have been previously identified and characterized. The others resulted from the analysis of published genomic sequence for *Drosophila melanogaster* and molecular studies undertaken in this study. The classification is based on domain and phylogenetic analysis of the predicted mRNAs for these myosins. In several cases the presence or absence of certain motifs was verified by RT-PCR (for detailed explanation of this, see text). The accession numbers show the most recent and complete sequence reports on the presented myosins. ESTs for all of the predicted genes have been identified and partially sequenced. In the case of *Mhcl* the presented mRNA sequences were combined and the three longest transcripts shown.

pression patterns (Bernstein and Milligan 1997; Zhang and Bernstein 2001). *Drosophila* has a second myosin II gene that encodes a cytoplasmic nonmuscle myosin (Kiehart and Feghali 1986; Mansfield et al. 1996). The founding member of class III myosins was discovered in *Drosophila* (Montell and Rubin 1988). The *ninaC-III* gene encodes two isoforms resulting from alternative RNA splicing. These differ in the composition of their C-terminal tails and show differential expression patterns (Porter et al. 1992; Li, Porter, and Montell 1998). A single myosin V gene, with at least two different splice forms, was identified in *Drosophila* (Bonafe and Sellers 1998; MacIver et al. 1998). Kellerman and Miller (1992) cloned a novel unconventional myosin from *Drosophila*—the first member of class VI myosins. The gene produces multiple protein isoforms, which are present throughout

Drosophila development (Kellerman and Miller 1992; Mermall and Miller 1995; Deng, Leaper, and Bownes 1999; Hicks et al. 1999). *Drosophila* myosin VIIA was the first member of this class to be described (Cheney, Riley, and Mooseker 1993; Kiehart et al. 1998).

While this manuscript was in preparation, another paper dealing with *Drosophila* myosins was published (Yamashita, Sellers, and Anderson 2000). The data from both research groups is complementary, with our manuscript focusing in detail on the molecular analysis and domain structure of the myosins.

Materials and Methods

Analysis of the *Drosophila* Genome—Analysis and Manipulation of Sequences

We have used the completed *D. melanogaster* Genome Project to determine the number of myosin-encoding genes in this species and to classify them. The new myosin genes were identified by comparing the *Drosophila* genome sequence with the conserved head (equivalent to amino acids 88–780) of the chicken skeletal myosin II. Comparison with the GenBank Data Base was done using the BLASTP algorithm of the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/BLAST/), the Pôle Bio-Informatique Lyonnaism server (http://pbil.univ-lyon1.fr/BLAST/blast.html), and The Berkeley Drosophila Genome Project BLAST (http://www.fruitfly.org/blast/). This search retrieved the eight previously described myosin genes as well as three new genes: *Myo 28B1* (AAF52536), *Myo10A* (AAF47980), and *Mhcl* (or *Myo89B*) (AAF55271). Subsequent detailed analysis was done using the conserved head or specific tails from all known classes of myosins to search the translated *Drosophila* genome sequence (TBLASTN on the NCBI server). These searches identified two other sequences highly similar to myosins: *Myo29D* (AAF52683, AF454348) and *Myo95E* (AAF56246, AF454350).

Domain Analysis—Multiple Sequence Alignment

The domain structure was predicted with the Simple Modular Architecture Research Tool (SMART) server at http://smart.embl-heidelberg.de/ (Schultz et al. 1998), the Pfam HMM database at http://pfam.wustl.edu/hmmsearch.shtml or http://www.sanger.ac.uk/Software/Pfam/search.shtml (The Sanger Centre), and the ProFile Scan Server of ISREC (Swiss Institute for Experimental Cancer Research) at http://www.isrec.isbsib.ch/software/PFSCAN_form.html. Alignments of the detected domains were performed with CLUSTAL W (Thompson, Higgins, and Gibson 1994) available from the Gene Jockey II software package distributed by Biosoft or from the WEB-based package at http://www2.ebi.ac.uk/clustalw/. Subsequently, the sequences were run on WEB-based BoxShade Server (http://ludwig-sun1.unil.ch:8080/software/BOX_form.html) and manually adjusted in Microsoft Word 98.

Coiled-coil regions were predicted with the Paircoil program developed by Berger and coworkers (Berger et al. 1995) at http://nightingale.lcs.mit.edu/cgi-bin/score.

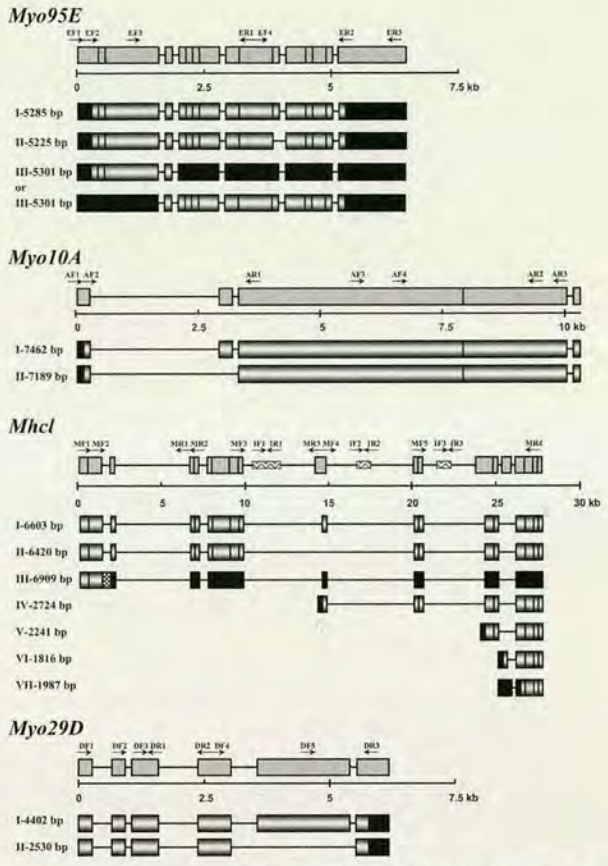


FIG. 3.—Molecular organization of *Myo95E, Myo10A, Mhcl,* and *Myo29D* myosin genes. The exon structure of the myosin genes is represented by gray boxes. Arrows indicate the primers used in the RT-PCR analysis of the gene structure. The length of the predicted mRNAs is noted on the left side of the transcript maps. Shaded boxes represent ORFs; solid boxes represent untranslated sequences (5′UTR or 3′UTR); dotted boxes represent the intron sequence in the spliced mRNA; wave pattern boxes represent exons from genes located within the myosin gene. The gene depicted by IF1 and IR1 primers is at least a three-exon gene, but for simplicity it is presented with a single box here (for more information see text).

RNA Extraction and RT-PCR

Total RNA from larvae and adult flies was isolated using RNAeasy-Total RNA Kit (QIAGEN, Cat #: 74104). The RNA (4 μg) was primed with Oligo-p$(dT)_{18}$ and reverse transcribed by Superscript™ II (GIBCO BRL), following the supplier's protocol. For subsequent DNA amplification, 2 μl out of 20 μl of the first-strand reaction mix was used. The PCRs were carried out with appropriate pairs of forward and reverse primers. The position of the primers is indicated in figure 3. Primers for *Myo95E:* EF1 (TGT TGC TCG CCA GCA CAT G), EF2 (ATG GAG CAG GAA ATC GGC AC), EF3 (CAT CAG CGG CCT TCC TGA AT), EF4 (GAA GTA CAT AAG CGC TGC CT), ER1 (AGG CAG CGC TTA TGT ACT TC), ER2 (ACA ATT ATC TCC ATG CGG TTC G), ER3 (ACG TAG ATG CCT GAA CTA TC). Primers for *Myo10A:* AF1 (GCA GCA ATG AAT CAA CCG GT), AF2 (TGA TCT GGT CTG GTT CGA TC), AF3 (TCA GTG TCC AGA GGC ATG TG), AF4 (TGG AGT GGC GTG CCT GGA), AR1 (CAT CTT GTA CGG ATT CAC

CG), AR2 (TCC GCA CGC GCA ACT TCC A), AR3 (CTT GCG GAA CTC CTG GAC A). Primers for *Mhcl:* MF1 (CAA CTT TAT GAA GAA GAG CGC), MF2 (AAG GCA GCT AGT GAT CAG GC), MF3 (TCG CAT AGG ACC CAG CCA G), MF4 (ATG TGG TCG GAT AAA AGT GCA), MF5 (GCT CTC AGA TCG CAT TAT ACA G), MR1 (AGC TCG CAG ATG TCC TCG A), MR2 (ACA TGG AGA CAA CCT TCT CG), MR3 (TAT ACG GAC GCA GGC GAT AG), MR4 (TCT TCC AGA TCA CTG ATA GAG). Primers for *Myo29D:* DF1 (ATC CGC ACA ACA TTC TGC AC), DF2 (ATG CAT CTT CAT CCA ACG AG), DF3 (AAT CAC AGC TTC AGC CAC AC), DF4 (GAG ACT GAT GCC TTC AAG CAC), DF5 (ACT TTG TGC GCT GCA TCC G), DR1 (GTC CCG ACA AGT GGA TCA G), DR2 (GTG CTT GAA GGC ATC AGT CTC), DR3 (AGG AAG AGT TGA ACA GAT GGA). Primers for genes in the *Mhcl* gene: IF1 (TTA CCT CCA TAA ACC TGC GG), IF2 (GTC GCC GAG CCC GAA GAG), IF3 (AAC GTC GCG TTC GCA AGA GG), IR1 (AAC GAT TCG GAG GTG CAC G), IR2 (CTA GCT CTG CGA AGA TCT CA), IR3 (CGT TCA TGG CTG CTA GTA CG). QIAGEN Taq and Stratagene Pfu Turbo® Polymerase in a 3:1 ratio were used in the PCR reactions. The QIAGEN PCR protocol designed to work with Q-Solution was followed. The PCR reactions were carried out as follows: one cycle at 94°C for 4 min; 35 cycles, step one—94°C for 40 s, step two—60°C for 40 s, step three—72°C for 1.8 min per expected kilobase pair of the PCR product; and one cycle at 72°C for 10 min. The obtained PCR products were isolated from Tris-acetate/Ethylenediaminetetra-acetate (TAE) gels, purified, and sequenced on a 373A automated DNA sequencer (ABI).

Results

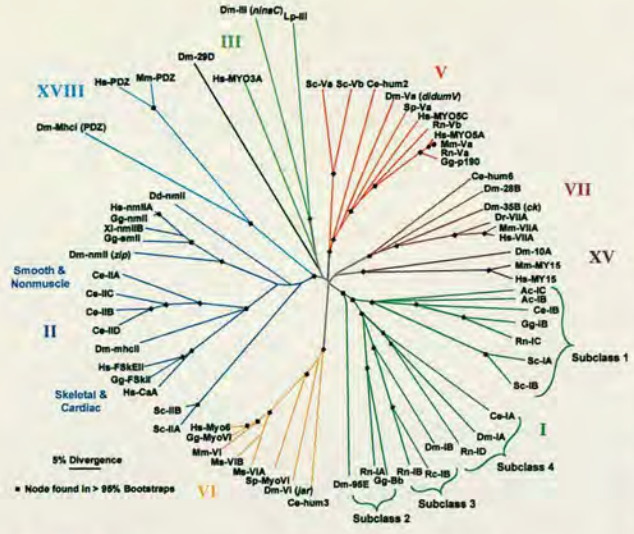
The BLASTP and BLASTN searches with the conservative chicken skeletal myosin II head, against the completed *Drosophila* genome sequence, retrieved the previously identified genes and five new myosin genes (fig. 1). We selected a limit of 30%–40% identity for a myosin to fall into a given myosin class, and 25%–30% was considered as the lower threshold for a protein to be classified as a myosin. The new myosins (fig. 1*B*) were named according to their chromosome position (fig. 2). Interestingly, half of the myosin genes are located between polytene bands 27F–36A, thus forming a myosin hot spot on the left arm of the second chromosome.

In the cases of *Myo 95E, Myo10A, Mhcl,* and *Myo29D,* we did a detailed analysis of the molecular structure of the genes and the transcripts they produced. The main reason for this was that the predicted sequences for these genes encode proteins that produced low homology scores to other myosins. They were obviously myosins, containing all the conserved sequences and structural parts defining them as myosins and at the same time showing no more than 15%–29% identity to other myosins. This implied either incorrectly predicted genes or incorrectly predicted splicing of the transcripts. Open reading frames (ORF), 5′untranslated regions (UTRs), 3′UTRs, and the presence or absence of given motifs were tested by RT-PCR and subsequent sequencing of the products obtained. In these experiments we employed an ovarian Uni-ZAP XR® library produced in our lab, along with cDNAs produced by reverse transcription of RNAs from larvae and adult flies (see *Materials and Methods*). As a result, we determined a number of new myosin sequences and submitted them to the MEDLINE Database. The accession numbers for these are—*Myo95E:* AF454350, AF454351, and AF454352; *Myo10A:* AF454346 (presents a part of the first three exons including the 5′UTR); *Mhcl:* AF454347 (presents a part of the first four exons including the sequence encoding the PDZ domain); *Myo29D:* AF454348 and AF454349. *Myo28B* was not subjected to detailed analysis because it was found to be almost identical to the other myosin VII (crinkled) from *Drosophila* at both the DNA and protein levels.

To examine the evolutionary relationships between members of the myosin family in *Drosophila* and other phylogenetically diverse species, we used two different phylogenetic methods. We applied Distant-matrix and Maximum-Parsimony methods (PROTDIST and PROTPARS from the PHYLIP package) to compare the conserved head domains. These methods were chosen because they tend to outperform other methods (i.e., lower variance), such as the Maximum Likelihood, when dealing with large data sets. The two programs produced trees with similar topology (see the unrooted consensus tree in fig. 4). Multiple sequence alignments were performed with CLUSTAL W without corrections for gaps or multiple substitutions. Excluding the positions with gaps would have omitted a significant proportion of the data, a problem that occurs when large amounts of input sequences are dealt with. CLUSTAL W (GCG software package) is provided by the Human Genome Mapping Project Resource Centre, Cambridge, at http://www.hgmp.mrc.ac.uk/ (Thompson, Higgins, and Gibson 1994). The reliability of the tree structure was checked by bootstrapping (1,000 trials) and reordering the alignments randomly (bootstrapping was performed with SEQBOOT from the PHYLIP package). The tests produced trees with similar branching order. A consensus tree was produced by the CONSENSE program of the PHYLIP package and graphically drawn with the TREEVIEW program (Page 1996), and was then transferred to and manipulated with PowerPoint.

The protein sequences for the new *Drosophila* myosins are theoretical predictions. Myosins are large multiexon genes and are difficult to assemble with 100% accuracy from sequence data. There are also various isoforms of some myosins, which can lead to some misalignments; hence, it is unlikely that the tree shown perfectly reflects the evolution of the *Drosophila* myosins.

It is possible that some of the new *Drosophila* myosins could be pseudogenes. However, we found that the probability for this was low. Pseudogenes generally lack introns and are not transcribed into mRNA. We identified expressed sequence tags (ESTs) for all the myosin genes we predicted (the accession numbers for these are given in table 2), which confirmed their in vivo expression. The new myosins are described in detail subsequently.



Abbrevation full terms:

Ac	*Acanthamoeba castellanii*	Lp	*Limulus polyphemus*	Bb	Brush border myosin I
Ce	*Caenorhabdites elegans*	Mm	*Mus musculus*	Ca	Cardiac alpha myosin II
Dd	*Dictyostellum discoideum*	Ms	*Morone saxatilis*	FSk	Fast skeletal muscle myosin II
Dm	*Drosophila melanogaster*	Rc	*Rana catesbeiana*	mhc	Myosin heavy chain
Gg	*Gallus gallus*	Rn	*Rattus norvegicus*	nm	Non-muscle myosin II
Hs	*Homo sapiens*	Sc	*Saccharomyces cerevisiae*	PDZ	PDZ containing myosin XVIII
		Xl	*Xenopus laevis*	sm	Smooth muscle myosin II

FIG. 4.—Phylogenetic analysis of the myosin superfamily in *D. melanogaster*. The head domains of *Drosophila* myosins (equivalent to amino acids 88–780 of chicken skeletal myosin) were compared with selected myosins representing phylogenetically diverse organisms. The alignment of 67 myosins (performed with the CLUSTAL W package) was used to generate data for an unrooted tree to scale based on sequence divergence (for details see text). Each myosin class is defined by the first node represented in >95% of the bootstrap trials (starting from the tree center) and is color coded. The accession numbers for the myosins used in the alignment are given subsequently. These are produced for the protein sequences, except in the case of Dm95E (*Myo95E* gene) and Dm29D (*Myo29D* gene), where the cDNA accession numbers are given. Myosin class I: Ac-IB (AAA27708), Ac-ICy (AAA27707), ScMyo3-IA (AAB34124), ScMyo5-IB (CAA89745), Cehum5-IA (CAA53244), CehumI-IB,1 Dm31D- IA (AAA19590), Dm61F-IB (AAA19591), Dm95E (AF454350), Rc-IB (AAA57192), Gg-Bb (AAB38373), Gg-IB (CAA49850), Rnmyr1-IA (CAA48287), Rnmyr2-IB (CAA52807), Rnmyr3-IC (CAA52815), Rnmyr4-ID (CAA50871). Myosin class II: ScMYO1-IIA (CAA37894), ScMYS1-IIB (P08964), Dd-nmII (AAA33227), Ce myo3-IIA (CAA30856), Cemyo4-IIB (AAA28124), Cemyo2-IIC (CAA30855), Cemyo1-IID (CAA30854), Dm-mhcII (AAA28686), Dmzip-nmII (AAF47311), Xl-nmIIb (AAA49915), Gg-FSkII (P13538), Gg-nmII (AAA48974), Gg-smII (CAA29793), Hs-FSkEII (CAA32167), Hs-CaA (BAA00791), Hs nmIIA (AAA59888). Myosin class III: Lp-III (AAC16332), Dm-NinaC-III (AAA28718), HsMYO3A (AAF70861). Myosin class V: ScMyo2p-Va (AAA34810), ScMyo4p-Vb (AAC37409), Cehum2 (AAA97926), SpMyo-Va (AAF78910), Dmdidum-Va (AAC99496), Ggp190-Va (CAA77782), Mmdilute-Va (CAA40651), Rnmyo5-Va (BAA88350), Rnmyr6-Vb (AAB38840), HsMYO5A (AAD00702), HsMYO5C (AAF78783). Myosin class VI: SpMyo-VI (AAF72176), Cehum3 (AAC67447), Dmjar-VI (CAA47462), Ms-VIA (AAD52005), Ms-VIB (AAD52006), GgMyoVI (CAB96536), MmWaltzerVI (AAB00194), HsMYO6 (AAC51654). Myosin class VII: Cehum6-VII (AAB37988), Dmck-IIA (AAF53435), Dm28B-VIIB (AAF52536), DrMyo-VIIA (CAC05418), Mmshaker-VIIA (AAB40708), HsUsherIb-VIIA (AAB03679). Myosin class XV: Dm10A-XV (AAF47980), Mm-Myo15 (AAF05904), HsMYO15 (AAF05903). Myosin class XVIII: Dm89B-PDZ (AAF55271), Mm-PDZ (BAA93660), Hs-PDZ (BAA13206).2 Myosin class ?: Dm29D (AF454348).

Myo95E (Myosin IC)

This myosin was not found during the initial searches of the fly genome. It was subsequently identified in the AE003746 genome scaffolding. The predicted sequence for this gene is unusually short, resulting in only a 59-aa protein (accession number AAF56246). Detailed analysis showed that the gene was not predicted correctly by the Genome Project. The sequence for *Myo95E* was manually assembled, taking into account the reported sequence for this gene along with the identified ESTs (table 2), as well as the homology of the translated genomic DNA to other myosins (preserving the exon-intron spacing). To test our theoretical predictions, we used two splice-site prediction programs: The Neural Network at http://www.fruitfly.org/seq_tools/splice.html and the GENSCAN Server at MIT—http://genes.mit.edu/GENSCAN.html. Using the artificially assembled sequence, we were able to design primers, amplify a PCR product, and hence sequence the cDNA for *Myo95E* (see *Materials and Methods*). The resulting sequence differed from the predicted sequence, showing the presence of an unusually long exon 3 and variations in exon 4. RT-PCR analysis showed that the gene produces at least three different transcripts expressed during oogenesis, and larval and adult stages (fig. 3). Two of the transcripts (I-AF454350 and II-AF454351), which represent 5%–10% of the total amount of *Myo95E* transcripts, translate into two protein isoforms of 1,278 and 1,258 aa, respectively. Transcript I (5,285 bp) comprises all 16 exons of the gene. Transcript II (5,225 bp) lacks exon 12. The major third transcript (III-AF454352) has a longer exon 4 (a downstream extension of 16 bp, GTG CAC ATT ACC CAT T). This shifts the ORF and produces a stop codon TGA in exon 5. The third transcript (5301 bp) encodes two putative truncated proteins. The first is a 464 aa sequence containing only the GESGAGKT conserved region from the P-loop of the myosin head domain. It does not contain the Switch-1 (NxxSSR) and Switch-2 (DxxGxE) regions which together with the P-loop have been implicated as having a role in the hydrolysis of ATP. The second is an 864 aa protein containing a large part of the head domain, two IQ motifs and a myosin tail domain. The head of this myosin form contains only the Switch-2 (DFYGFE) conserved sequence, which prevents it from converting ATP, probably rendering the form inactive. Homology searches with the protein sequence for the head domain showed 33% identity (53% similarity) to vertebrate brush border class I myosins. Three ESTs were found for *Myo95E*. A search with them revealed up to 37% identity to vertebrate brush border myosins and less to other myosin classes. Analysis of Myosin 95E with domain-scanning programs confirmed its structural similarity to class I myosins (fig. 1*B*). It was found to contain two IQ domains (928–974 aa), the second being poorly conserved (fig. 5),

←

[1] This is a joint sequence from the first 72 amino acids of T21544 and from the 11th amino acid to the end of AAA97925.

[2] The sequence for the human PDZ-containing myosin is truncated. This does not seem to affect its evolutionary position in the phylogenetic tree.

FIG. 5.—Alignment of the IQ motifs for the five newly identified myosins from *D. melanogaster*: Myosin IC (*Myo95E*), Myosin VIIB (*Myo28B*), Myosin XV (*Myo10A*), Myosin 89B (*Mhcl*), and Myosin 29D (*Myo29D*). The consensus sequence is shown below the alignment, and the highly conservative positions are shaded. Residues in bold indicate similarity to the consensus sequence. Myosin IC has two different variants (A and B) of its second IQ motif, a result of alternative splicing.

and a Basic Tail domain (974–1,278 aa for isoform I and 974–1,258 aa for isoform II) (fig. 1*B* and table 1). The latter is thought to be involved in membrane binding. Recent studies have shown that it can also bind to actin filaments (Lee et al. 1999; Liu, Brzeska, and Korn 2000). This changes the number of class I myosins in *Drosophila* to three, hence Myosin 95E was renamed Myosin IC. The *Drosophila* Myosin IC does differ from other myosins of class I. The unusually long exon 3 results in a 281-aa insertion into the head domain. This insertion contains a partial AAA domain, a conserved region that contains an ATP-binding site. So far, no other myosins from class I have been identified which contain such an insertion.

Myo28B (Myosin VIIB)

The amino acid sequence of the Myosin 28B head showed 61% identity (74% similarity) to *ck-Drosophila* myosin VIIA. It also exhibited a very high identity of 58% (72% similarity) to Myosin VIIA from zebrafish and to other class VII myosins. Analysis of Myosin 28B revealed that it has four IQ motifs (753–845 aa), the third being poorly conserved (figs. 1*B* and 5). Two Myosin Tail Homology 4 (MyTH4) (1,070–1,246 and 1,681–1,826 aa), two FERM (1,246–1,454 and 1,826–2,039 aa), and one SH3 (Src homology 3) (1,561–1,626 aa) domains were identified (fig. 6*A* and *B*). The function of the MyTH4 domains is unknown. The FERM domain (the name stands for Band 4.1, ezrin, radixin, moesin-homology) is believed to be involved in linking cytoskeletal proteins to the membrane as well as in dimerization. Talin, merlin, and philopodin are other major members of the FERM superfamily. These deserve mentioning, especially the Talin, because the FERM domain exhibits the highest homology to the FERM domain in Talins (fig. 7) and less homology to FERM domains from other members of the FERM family. The SH3 domain has been identified in many proteins involved in signal transduction. It is believed that the SH3 domains mediate protein-protein interactions by binding to proline-rich domains. Other myosins, such as IV, X, and XV, also contain this

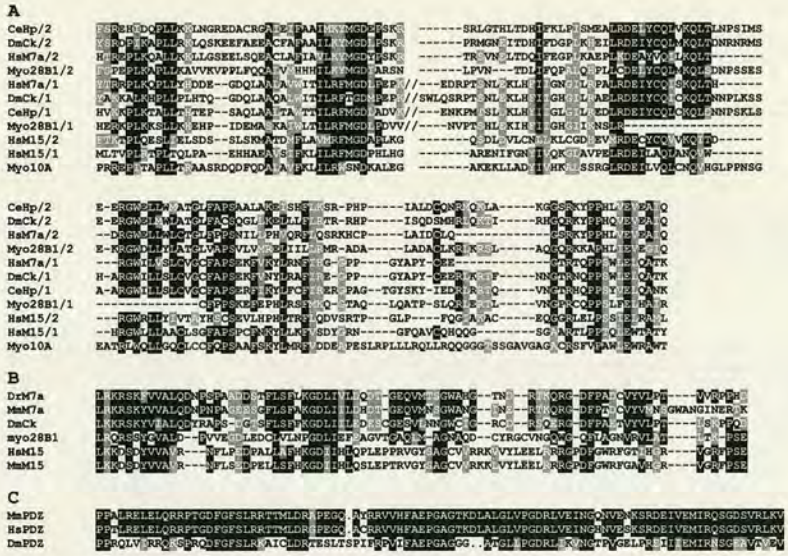


FIG. 6.—CLUSTAL W alignment of the predicted (*A*) MyTH4, (*B*) SH3, and (*C*) PDZ domains in the tails of the newly identified *Drosophila* myosins. The sequences were shaded on the BoxShade server and manipulated with Microsoft Word 98. The parameters set on BoxShade are as follows—output format: RTF New; fraction of sequences: 0.5; input sequence format: MSF or multisequence format (this is the output format of CLUSTAL W). Identical amino acids are shaded in black, similar residues are shaded in gray. Myosin VII has two MyTH4 domains. The first MyTH4 domain in HsM7a, DmCk, CeHp, and Myo28B1 contain an insert of 61–88 aa (regions 1057–1144, 1048–1135, 971–1055, and 1107–1168, respectively) that was not included in the alignment. Accession numbers and abbreviation full terms: Ce, *Caenorhabditis elegans* (Myosin VII, T25888); Hp, hypothetical protein; Dm, *Drosophila melanogaster* (ck, VIIA, AAF44915; Myo28B, AAF52536; PDZ-Myosin, AE003711); ck, crinkled; Hs, *Homo sapiens* (M7A, Q13402; Myosin XV, A59266; PDZ-Myosin, AL080245 and Z98949); Dr, *Danio rerio* (Myosin VIIA, CAC05418); Mm, *Mus musculus* (Myosin VIIA, P97479; Myosin XV, A59295; PDZ-Myosin, BAA93660).

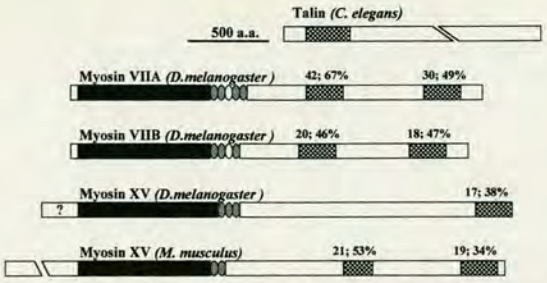


FIG. 7.—Sequence comparison of the FERM domains found in *Drosophila* Myosin VII and XV. These were aligned separately to the Talin domain from *C. elegans* Talin protein. Multiple alignment was not appropriate in this case because the similarity between FERM 1 and FERM 2 to the Talin domain (from the Talin protein) is very low and results in unordered alignment. The comparison was performed by CLUSTAL W (http://www2.ebi.ac.uk/clustalw/). In addition, the FERM domains from *Mus musculus* Myosin XV were included in the diagram. In general, FERM1 motifs from Myosin VII show ~12%–21% similarity to proteins from the FERM family and ~16% to the band 4.1 protein. FERM2 domains (Myosin VII) show a higher similarity of ~17%–18% to band 4.1 protein and ~18%–20% to proteins from the FERM family. FERM domains from Myosin XV show less similarity of 9%–17% to FERM proteins and ~14% to the band 4.1 protein. Accession numbers are—*C. elegans* (Talin): AAA74747; *D. melanogaster* (Myosin VIIA): CAC05418, (Myosin VIIB): AAF52536, (Myosin 10A-XV): AAF47980; *M. musculus* (myosin XV): AAF05904.

motif. A short coiled-coil domain (849–908 aa) was predicted by the Paircoil program (fig. 8).

Myo10A (Myosin XV)

An RT-PCR analysis was used to verify the exon composition and the exon length of *Myo10A* transcripts. It was found that there are at least two transcripts, which are expressed from the larval stage onward (fig. 3). The longer transcript (7,462 bp) consists of all the previously known five exons and represents no more than 5% of the total amount of mRNA for this myosin. The shorter transcript (7,189 bp) lacks exon 2 and is expressed abundantly. The two transcripts translate into two protein isoforms of 2,424 and 2,333 aa, respectively. A BLAST search with the conserved head domain showed significant 47% identity (64% similarity) to mouse and human myosin XV. *Myo10A* is also related to human and mouse myosin VIIA with an identity of 42% (59% similarity). A specific N-terminal domain was identified in Myosin 10A (1–149 aa). The latter showed no similarity to the characteristic N-terminal domain found in other class XV myosins. The shorter protein isoform lacks this N-terminal domain. In the neck region three IQ domains (841–910 aa) were identified (fig. 5). Immediately after the IQ motifs there is a short coiled-coil region (919–946 aa) (fig. 8). Analysis of the tail revealed the presence of one MyTH4 domain (1,014–1,173 aa) (fig. 6), a glutamine-proline-alanine (QPA)-rich domain, a proline-rich domain (the borders of QPA and the proline-rich domain were not clearly defined), a short transmembrane motif (2,194 aa) (fig. 1*B*), and an FERM domain (2,220–2,424 aa). The latter showed a very high identity of 40% (59% similarity) to the first and 13% identity (30% similarity) to the second FERM domain from mouse myosin XV and a limited similarity of 38% (17% identity) to Talin itself (fig. 7).

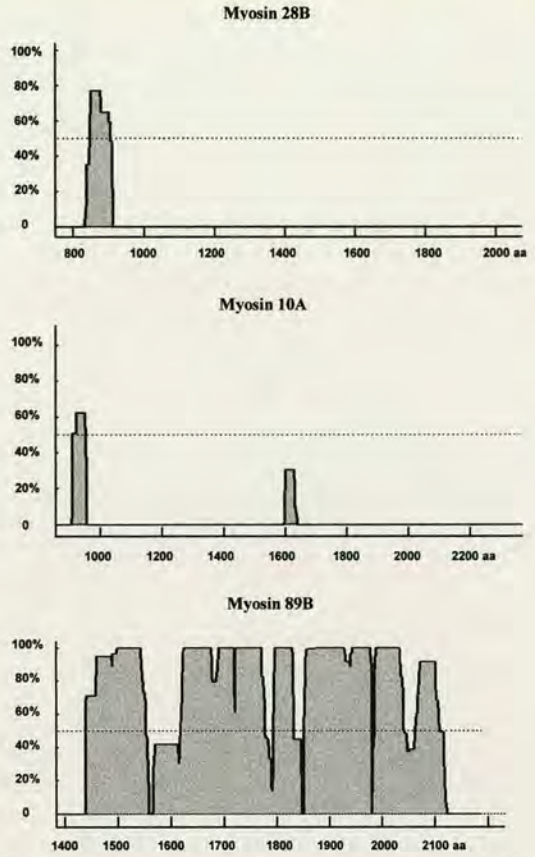


FIG. 8.—Predicted coiled-coil structures in the tails of the newly identified *Drosophila* myosins. Only three of the five myosins are likely to form coiled-coils, these being myosin 28B-VIIB, myosin 10A-XV, and Mhcl-PDZ. Both myosin 28B-VIIB and 10A-XV produced scores close to the lower threshold. Whether they form coiled-coils or not can only be determined experimentally.

Mhcl (Myosin XVIII)

A partial sequence of this gene has been submitted to the GenBank Data Base by Biru (1999). It shows high similarity to the Mhc type II of myosins and was termed Myosin heavy chain–like. Subsequently, the full length for this myosin was published by Celera (Adams et al. 2000). The RT-PCR analysis we have undertaken shows that *Mhcl* has a very complex structure and expresses multiple transcripts throughout the *Drosophila* life cycle (fig. 3). Data from the available ESTs for *Mhcl*, PCR analysis, and a Promoter-predicting program (http://www.fruitfly.org/seq_tools/promoter.html) suggest that the seven identified transcripts are transcribed from four putative promoters. Transcripts I (6,603 bp) and III (6,909 bp) are highly expressed. The other five identified transcripts are expressed at comparatively lower levels. Transcripts I and III are composed of 18 exons (the gene comprises 19 exons), lacking exon 15. The difference between them is that transcript III has the second intron spliced out, which introduces a stop codon in the ORF. One unusual feature of *Mhcl* is the fact that it contains three other genes within it (fig. 3). The exons of these genes are not found in combination with any of the exons of the myosin gene. Importantly, there are ESTs for each of the three genes. One of the genes has been previously

identified as phosphatidylserine-specific phospholipase A1 (CG4979). It is located in intron 9, and its reading frame is in the opposite direction to that of the *Mhcl* reading frame (EST: GH15759). The other two genes have ORFs in the same direction as the main (*Mhcl*) gene. The gene in intron 10 (ESTs: RE41368, RE44374) is a novel gene and does not show any significant homology to previously characterized genes. The gene in intron 12 (ESTs: LP08646, LP05315) shows a limited identity of 35% (46% similarity) to bovine synaptojanine 1 protein (synaptic inositol-1,4,5-trisphosphate 5-phosphatase 1). The *Mhcl* transcripts translate into a set of protein isoforms—I: 2,200 aa; II: 2,139 aa; III: 439 aa; IV: 899 aa; V: 730 aa; VI: 512 aa; and VII: 479 aa. Only the isoforms produced by transcripts I and II contain the conserved motor domain. BLAST analysis showed that Myosin heavy chain–like is most closely related to mouse PDZ-Myosin (41% identity) and human PDZ-Myosin (40% identity). It showed a limited identity of 28%–29% to smooth muscle and nonmuscle myosins and less than 25% to skeletal muscle myosin. BLAST searches revealed that the tail domain shows the highest similarity, 32%, to PDZ-containing myosins and less than 25% similarity to other types of myosins. Interestingly, this myosin was not predicted to contain a PDZ domain. We used the DNA sequence encoding the mouse PDZ domain to search the genomic sequence (AE003711) surrounding *Mhcl* (approximately 10 kb in each direction). This detected a PDZ domain 4,810 bp upstream of the predicted start for the *Mhcl* mRNA (the domain was found in the borders of the predicted genomic DNA for this gene but has not been previously included in its ORF). The presence of the PDZ domain was verified by RT-PCR, and the sequence was submitted to the NCBI database (AF454347). A sequence alignment of the PDZ domains from Drosophila, mouse, and human is given in figure 6. Interestingly, transcript III translates into a truncated protein containing only this PDZ domain (349–429 aa). PDZ domains are known to bind C-terminal or internal (non–C-terminal) polypeptides. Two IQ domains were found in Mhcl (1,379–1,428 aa) (figs. 1*B* and 5), although the second differed slightly from the consensus sequence. The Paircoil program predicted two coiled-coil domains at 1,439–1,549 and 1,616–2,121 aa (fig. 8).

Myo29D

Database searches suggested that the first three exons of *Myo29D* are probably part of a gene adjacent to the myosin. Subsequent RT-PCR showed that they splice together with the other three myosin exons to produce at least two different transcripts. Transcript I (4,402 bp) contains all the six exons, whereas transcript II (2,530 bp) lacks exon 5. Transcripts I and II are present during all the stages of *Drosophila* development, with transcript I being expressed at the higher levels. Transcript II produces a truncated form of the protein, lacking most of the motor head domain. Analysis of the conserved head domain revealed that Myosin 29D is not similar to any of the known classes of myosins. It showed 29% identity (45% similarity) to myosin VII, X, and V from different species, which is sufficient for it to be considered as a myosin but not enough to be related to a given class. A search with the available ESTs for this myosin showed 36%–39% identity to myosin V from different species and 31%–36% to vertebrate Myosin heavy chains (myosin II). Motif search programs detected in Myosin 29D a specific N-terminal extension (1–338 aa), two IQ motifs (1,089–1,136 aa), the second being poorly conserved, and a short transmembrane motif in the tail (1,144–1,279 aa) (fig. 1*B*).

The detailed search of the translated *Drosophila* genome sequence produced several low-score hits. Close analysis of these (searches up and downstream of the respective sequences) revealed that they are not true myosins. Despite exhibiting certain similarities to given parts of the myosins, none of them showed high homology to a larger part of the head domain. AE003112 showed similarity to the highly conserved region (GESGAGKT) from the P-loop of the head domain. Its polytene chromosome location remains undetermined. AE003614 (CG11199) showed homology to the myosin tail, containing coiled-coil and FERM domains. It is located on the chromosome arm 2L (27E), next to the *ninaC* gene. AE003495 (CG12047) also showed similarity to the myosin tail, containing coiled-coil and FERM domains. This is located on the X chromosome (12E1), close to the newly identified *Myo10A*.

The BLAST search retrieved one more myosin-like gene *CG15831* in the AE002795 genome scaffolding. This is a single-exon gene, 219 bp long. The chromosome position of AE002795 has not been determined previously. The program showed similarity to a highly conservative part of the myosin head domain (**LGV LDIFGF**ENFSH**NSFEQ**LCI**NYT**N**EKLHKFFN**H). We found that the DNA surrounding the gene shows no similarity to the myosin genes (at the DNA or protein level). This suggests that AE002795 is either incorrectly assembled in the genome or that the gene *CG15831* is a partial duplication of a myosin gene (this might be either a part of *Mhc* [muscle myosin II] or *didum* [Myosin V], which produced the highest score for this sequence, 59% and 54% identity, respectively).

Discussion

The myosin family has grown significantly in the past decade to encompass more than 177 myosins these days. All these myosins, with a few exceptions, fall into 18 classes. Myosins are expressed in both prokaryotes (though these have not been well studied) and eukaryotes.

The annotation of the genome sequences for *S. cerevisiae*, *Dictyostelium discoideum*, *Arabidopsis thaliana*, *C. elegans*, *Drosophila melanogaster*, and *Homo sapiens* made it possible to identify the complete sets of myosin genes in these organisms. It appears that all eukaryotes have an essential set of three myosin genes, these being from classes I, II, and V, as well as a number of species-specific myosins. In *Drosophila* 13 different myosin genes were identified. The fruitfly has three I, two II (one encoding for muscle and one for nonmuscle myosins), one III, one V, one VI, two VII, one XV, and one XVIII myosin

genes, and one yet to be classified myosin gene. These classes of myosins have been found in a wide range of invertebrate and vertebrate animals. It has been shown that they have a role in a variety of cell functions, including membrane trafficking, signal transduction, and maintenance of the cell architecture. Several new myosin genes were identified in the fruitfly. The genome data provided cDNA sequences for four of them (*Myo28B, Myo10A, Mhcl,* and *Myo29D*). The sequences and the ORF for these myosins were verified by RT-PCR and sequencing. In addition, a fifth myosin was found (*Myo95E*). Using the fruitfly genomic sequence, we were able to predict and subsequently isolate and sequence this gene.

Only two of the five newly identified genes in *Drosophila* fell directly into previously known classes, these being Myosin VIIB (*Myo28B*) and Myosin heavy chain–like (*Mhcl*). Myosin VIIB is from the well-studied class VII myosins with a role in the membrane trafficking and stereocilia function.

Mhcl is a member of the recently defined class XVIII of myosins, which consists of only two other members, mouse and human PDZ-containing myosins. The PDZ domain, also called DHR (Dlg homologous region), is known to bind either C-terminal or internal (non–C-terminal) polypeptides. These domains have been identified in a broad range of signaling proteins from bacteria, yeast, plants, insects, and vertebrates. PDZ domains have been implicated in targeting signaling molecules to submembranous sites.

Myosin IC (*Myo95E*) is the third member of class I myosins in *Drosophila*. It contains the Basic Tail domain (TH1) specific to all class I myosins. It is thought that TH1 binds to acidic phospholipids and actin filaments. Interestingly, this myosin contains an additional N-terminal insertion, which is similar to the AAA motif, a conserved region of about 220 aa that contains an ATP-binding site. This domain is inserted in the region of the loop 1 of the motor domain, the other region associated with the hydrolysis site for ATP, and probably modulates its activity.

Myo10A is closely related to class XV myosins. This myosin has a short N-terminal extension, which differs from the N-terminal domain characteristic for vertebrate myosins of class XV. It also lacks the SH3 and the second MyTH4 domain found in other myosin XV tails. Instead, the *Drosophila* Myosin XV tail contains three addition motifs, a glycine-proline-glutamine–rich domain, a proline-rich domain, and a small transmembrane domain. The proline-rich sequences have been demonstrated to bind to SH3, a small 50-aa motif. SH3 domains have been identified in a wide variety of intracellular and membrane-associated proteins and are implicated in signal transduction, linking signals transmitted from the cell surface by protein tyrosine kinases to effector proteins located downstream on the hierarchical pathways.

Myosin 29D is a highly divergent member of the myosin superfamily. Presently, it forms a class of its own. This myosin contains an unusual N-terminal extension, which shows no homology to other proteins. It also has a small transmembrane domain in its tail rich in proline residues.

Drosophila melanogaster expresses many myosins genes. Apart from the essential myosins (classes I, II, and V), it also has myosins from classes III, VI, VII, XV, and XVIII, as well as a novel type of myosin. This new data should help to design experiments to investigate the roles of these newly identified myosins in the cell histology and development of *Drosophila*.

Acknowledgments

We are grateful to Neville Cobbe for help with the phylogenetic analysis programs, to Gillian Millburn from flybase for comments on the manuscript, and to Sheila Milne for help with the preparation of the manuscript. We are grateful to the Darwin Trust, MRC, BBSRC, and B'nei B'rith Scholarship Committee for their support.

LITERATURE CITED

ADAMS, M. D., S. E. CLINKER, R. A. GIBBS, G. M. RUBIN, and C. J. VENTER. 2000. The genome sequence of Drosophila melanogaster. Science **287**:2185–2195.

BAKER, J. P., and M. A. TITUS. 1997. A family of unconventional myosins from the nematode *Caenorhabditis elegans*. J. Mol. Biol. **272**:523–535.

———. 1998. Myosins: matching functions with motors. Curr. Opin. Cell Biol. **10**:80–86.

BERG, J. S., B. H. DERFLER, C. M. PENNISI, D. P. COREY, and R. E. CHENEY. 2000. Myosin-X, a novel myosin with pleckstrin homology domains, associates with regions of dynamic actin. J. Cell Sci. **113**:3439–3451.

BERG, J. S., B. C. POWELL, and R. E. CHENEY. 2001. A millennial myosin census. Mol. Biol. Cell **12**:780–794.

BERGER, B., D. B. WILSON, E. WOLF, T. TONCHEV, M. MILLA, and P. S. KIM. 1995. Predicting coiled coils by use of pairwise residue correlations. Proc. Natl. Acad. Sci. USA **92**: 8259–8263.

BERNSTEIN, S. I., and R. A. MILLIGAN. 1997. Fine tuning a molecular motor: the location of alternative domains in the *Drosophila* myosin head. J. Mol. Biol. **271**:1–6.

BIRU, W. G. 1999. Thesis, Attila Jozsef University, Szeged, Hungary. *Drosophila melanogaster* mRNA for myosin heavy chain, partial. GenBank/EMBL/DDBJ: AJ132656.

BONAFE, N., and J. R. SELLERS. 1998. Molecular characterization of myosin V from *Drosophila melanogaster*. J. Muscle Res. Cell Motil. **19**:129–141.

BROWN, S. S. 1997. Myosins in yeast. Curr. Opin. Cell Biol. **9**:44–48.

CHENEY, R. E., M. A. RILEY, and M. S. MOOSEKER. 1993. Phylogenetic analysis of the myosin superfamily. Cell Motil. Cytoskelet. **24**:215–223.

COPE, M., T. V. JAMIE, J. WHISSTOCK, I. RAYMENT, and J. KENDRICK-JONES. 1996. Conservation within the myosin motor domain: implications for structure and function. Structure **4**:969–987.

DENG, W.-M., K. LEAPER, and M. BOWNES. 1999. A targeted gene silencing technique shows that *Drosophila* myosin VI is required for egg chamber and imaginal disc morphogenesis. J. Cell Sci. **112**:3677–3690.

FURUSAWA, T., S. IKAWA, N. YANAI, and M. OBINATA. 2000. Isolation of a novel PDZ-containing myosin from hematopoietic supportive bone marrow stromal cell lines. Biochem. Biophys. Res. Commun. **270**:67–75.

GILBERT, S. P., and A. T. MACKEY. 2000. Kinetics: a tool to study molecular motors. Methods **22**:337–354.

HAMMER, J. A., and G. JUNG. 1996. The sequence of the dictyostelium myo J heavy chain gene predicts a novel, dimeric, unconventional myosin with a heavy chain molecular mass of 258 kDa. J. Biol. Chem. **271**:7120–7127.

HASSON, T., and M. S. MOOSEKER. 1995. Molecular motors, membrane movements and physiology: emerging roles for myosins. Curr. Opin. Cell Biol. **7**:587–594.

HASSON, T., J. F. SKOWRON, D. J. GILBERT et al. (15 co-authors). 1996. Mapping of unconventional myosins in mouse and human. Genomics **36**:431–439.

HASTINGS, G. A., and C. P. EMERSON JR. 1991. Myosin functional domains encoded by alternative exons are expressed in specific thoracic muscles of *Drosophila*. J. Cell Biol. **114**:263–276.

HICKS, J. L., W.-M. DENG, A. D. ROGAT, K. G. MILLER, and M. BOWNES. 1999. Class VI unconventional myosin is required for spermatogenesis in *Drosophila*. Mol. Biol. Cell **10**:4341–4353.

HODGE, T., and M. J. T. V. COPE. 2000. A myosin family tree. J. Cell Sci. **113**:3353–3354.

KAWASAKI, H., S. NAKAYAMA, and R. H. KRETSINGER. 1998. Classification and evolution of EF-hand proteins. Biometals **11**:277–295.

KELLERMAN, K. A., and K. G. MILLER. 1992. An unconventional myosin heavy chain gene from *Drosophila melanogaster*. J. Cell Biol. **119**:823–834.

KIEHART, D. P., and R. FEGHALI. 1986. Cytoplasmic myosin from *Drosophila melanogaster*. J. Cell Biol. **103**:1517–1525.

KIEHART, D. P., R. A. MONTAGUE, J. ROOTE, and M. ASHBURNER. 1998. Evidence that crinkled, mutations in which cause numerous defects in *Drosophila* morphogenesis, encodes a myosin VII. Mol. Biol. Cell Suppl. **9**:388a.

KORN, E. D. 2000. Coevolution of head, neck, and tail domains of myosin heavy chains. Proc. Natl. Acad. Sci. USA **97**: 12559–12564.

LANGFORD, G. M. 1995. Actin- and microtubule-dependent organelle motors: interrelationships between the two motility systems. Curr. Opin. Cell Biol. **7**:82–88.

LEE, W. L., E. M. OSTAP, H. G. ZOT, and T. D. POLLARD. 1999. Organization and ligand binding properties of the tail of Acanthamoeba myosin-IA. Identification of an actin-binding site in the basic (tail homology-1) domain. J. Biol. Chem. **274**(49):35159–35171.

LI, H. S., J. A. PORTER, and C. MONTELL. 1998. Requirement for the NINAC kinase/myosin for stable termination of the visual cascade. J. Neurosci. **18**:9601–9606.

LIU, X., H. BRZESKA, and E. D. KORN. 2000. Functional analysis of tail domains of Acanthamoeba myosin IC by characterization of truncation and deletion mutants. J. Biol. Chem. **275**(32):24886–24892.

MACIVER, B., A. MCCORMACK, R. SLEE, and M. BOWNES. 1998. Identification of an essential gene encoding a class-V unconventional myosin in *Drosophila melanogaster*. Eur. J. Biochem. **2573**:529–537.

MANSFIELD, S. G., D. Y. AL-SHIRAWI, A. S. KETCHUM, E. C. NEWBERN, and D. P. KIEHART. 1996. Molecular organization and alternative splicing in zipper, the gene that encodes the *Drosophila* nonmuscle myosin II heavy chain. J. Mol. Biol. **255**:98–109.

MERCER, J. A., P. K. SEPERACK, M. C. STROBEL, N. G. COPELAND, and N. A. JENKINS. 1991. Novel myosin heavy chain encoded by murine dilute coat colour locus. Nature **349**: 709–713.

MERMALL, V., and K. G. MILLER. 1995. The 95F unconventional myosin is required for proper organization of the *Drosophila* syncytial blastoderm. J. Cell Biol. **129**:1575–1588.

MERMALL, V., P. L. POST, and M. S. MOOSEKER. 1998. Unconventional myosins in cell movement, membrane traffic, and signal transduction. Science **279**:527–533.

MONTELL, C., and G. M. RUBIN. 1988. The *Drosophila ninaC* locus encodes two photoreceptor cell specific proteins with domains homologous to protein kinases and the myosin heavy chain head. Cell **52**:757–772.

MOOSEKER, M. S., and R. E. CHENEY. 1995. Unconventional myosins. Annu. Rev. Cell Dev. Biol. **11**:633–675.

MORGAN, N. S., M. B. HEINTZELMAN, and M. S. MOOSEKER. 1995. Characterization of myosin-IA and myosin-IB, two unconventional myosins associated with the *Drosophila* brush border cytoskeleton. Dev. Biol. **172**:51–71.

MORGAN, N. S., D. M. SKOVRONSKY, S. ARTAVANIS-TSAKONAS, and M. S. MOOSEKER. 1994. The molecular cloning and characterization of *Drosophila melanogaster* myosin-IA and myosin-IB. J. Mol. Biol. **239**:347–356.

OLIVER, T. N., J. S. BERG, and R. E. CHENEY. 1999. Tails of unconventional myosins. Cell Mol. Life Sci. **56**:243–257.

PAGE, R. D. M. 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. **12**:357–358.

PORTER, J. A., J. L. HICKS, D. S. WILLIAMS, and C. MONTELL. 1992. Differential localizations of and requirements for the two *DrosophilaninaC* kinase/myosins in photoreceptor cells. J. Cell Biol. **116**:683–693.

RHOADS, A. R., and F. FRIEDBERG. 1997. Sequence motifs for calmodulin recognition. FASEB J. **11**:331–340.

SCHULTZ, J., F. MILPETZ, P. BORK, and C. P. PONTING. 1998. SMART, a simple modular architecture research tool: identification of signaling domains. Proc. Natl. Acad. Sci. USA **95**:5857–5864.

SCHWARZ, E. C., H. GEISSLER, and T. A. SOLDATI. 1999. Potentially exhaustive screening strategy reveals two novel divergent myosins in Dictyostelium. Cell Biochem. Biophys. **30**:413–435.

SELLERS, J. R. 2000. Myosins: a diverse superfamily. B. B. A.—Mol. Cell Res. **1496**:3–22.

SOLDATI, T., H. GEISSLER, and E. C. SCHWARZ. 1999. How many is enough? Exploring the myosin repertoire in the model eukaryote *Dictyostelium discoideum*. Cell Biochem. Biophys. **30**:389–411.

THOMPSON, J. D., D. G. HIGGINS, and T. J. GIBSON. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673–4680.

TITUS, M. A. 1999. A class VII unconventional myosin is required for phagocytosis. Curr. Biol. **9**:1297–1303.

VOLKMANN, N., and D. HANEIN. 2000. Actomyosin: law and order in motility. Curr. Opin. Cell Biol. **12**:26–34.

WOLFE, K. H., and D. C. SHIELDS. 1997. Molecular evidence for an ancient duplication of the entire yeast genome. Nature **387**:708–713.

WU, X., G. JUNG, and J. A. HAMMER. 2000. Functions of unconventional myosins. Curr. Opin. Cell Biol. **12**:42–51.

YAMASHITA, R. A., J. R. SELLERS, and J. B. ANDERSON. 2000. Identification and analysis of the myosin superfamily in *Drosophila*: a database approach. J. Muscle Res. Cell Motil. **21**:491–505.

ZHANG, S., and S. I. BERNSTEIN. 2001. Spatially and temporally regulated expression of myosin heavy chain alternative exons during *Drosophila* embryogenesis. Mech. Dev. **101**:35–45.

ANTONY DEAN, reviewing editor

Accepted February 7, 2002